

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 3 日現在

機関番号：13302

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330434

研究課題名(和文) 進化言語学の方法による音楽の文脈自由規則の発見

研究課題名(英文) Detection of Context-free Grammar in Music with Evolutionary Linguistics

研究代表者

東条 敏 (Tojo, Satoshi)

北陸先端科学技術大学院大学・情報科学研究科・教授

研究者番号：90272989

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：音楽の和音進行のカデンツ規則は、自然言語の文脈自由文法のクラスに属する。本研究では楽曲から文脈自由規則を発見する試みとして、音楽理論GTTMに基づくタイムスパン木に係り受け関係を見出すこと、進化言語学のモデルILMによる文法発見の規則を効率化することの両面から研究を行った。まずタイムスパン木にはピッチ(音高)の概念を組み込み、それにより和音が発見できることから、和音連鎖列の妥当性をTPS理論から評価した。さらにタイムスパンの構成から楽曲間の距離を測る方法を考案し、この数値的距離と心理的類似性との関係を研究した。進化言語学の方法においては文脈自由規則の短縮化などを用いILMの効率化を図った。

研究成果の概要(英文)：There exist context-free grammar (CFG) rules in music, e.g., for sequences of chords as cadences. In this research, we have investigated the formalisms to detect such CFG rules in music from two different aspects. First, we relied on the Generative Theory of Tonal Music (GTTM) and found a long-distance dependency, that is a witness for being CFG, in its time-span trees. Second, we have tried a method on evolutionary linguistics called the iterated learning model (ILM), and improved the efficiency of the grammar detection. In time-span tree, we introduced the notion of pitch, and thus we could detect the most plausible sequence of chords, employing the theory of tonal pitch space (TPS). In addition, we have defined the distance between two time-span trees, by the sum of different time-spans, and then we have evaluated the correlation between this distance and our psychological similarity. As for ILM, we have invented a method of string clipping, and thus we could improve the efficiency.

研究分野：人工知能

キーワード：音楽理論 文脈自由文法 カデンツ 繰り返し学習モデル GTTM TPS

1. 研究開始当初の背景

(1) 情報処理の観点から: 音楽は作曲・演奏・鑑賞など、あるゆるフェーズにおいて人間の情動と関わるため、形式的な文字列を処理するシンボル操作計算の対象とはみなしにくい。実際、音楽情報処理では音響・音声処理および検索・推薦システムなど統計的処理による計算機処理が盛んである。一方で、楽譜をシンボル列と考えた上で、記号列変形処理を行う研究-例えば楽譜を musicXML など表現し簡素化・装飾・合成など編曲する処理-は隠れマルコフモデルなど正規文法による規範の域を出ないため、後述するように音楽の大域的な性質を捉えることはできない。シンボル化は人間の直観・主観を形式化する意味で、認知科学視点からも大変重要であるにも関わらず、音楽の記号的取扱いは個々の研究毎にアドホックな方法に依存しているのが実情である。

(2) 音楽学の観点から: 音楽には階層的な構造が存在するとされる考え方は20世紀当初より H. Schenker により論じられ広く支持されてきたが、その抽象化された全体像(ゲシュタルト)を取り出すための簡約(reduction)法は具体的に示されないままだった。F. Lehrdahl と R. Jackendoff による Generative Theory of Tonal Music (GTTM) はこの簡約の考えを推し進め、音楽には言語同様の生成規則があると考え、音楽全体が木構造を構成するとした理論を展開した。われわれはまたこの理論に基づいて音楽の解析システムを実装し、その応用を発表してきた。しかしながら GTTM においても遠隔の依存関係を直接反映することは困難である。

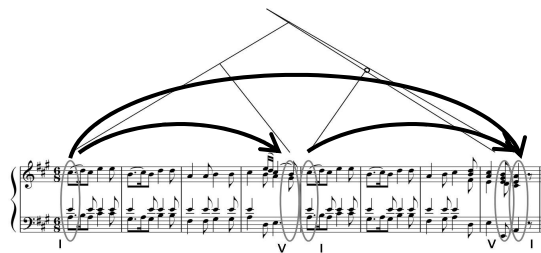


図1 モーツァルトK.331 冒頭の係り受け

図1では第4小節に限るとI(トニック)の和音のほうが優位であるが、半終止を表現するためには第6拍のV(ドミナント)が優位である。これはカデンツの意味で曲頭のIからの係り受けである。

(3) 言語学の観点から: 音楽を言語のように見立て文法を設定してパーサを作ろうという試みは T. Winograd を嚆矢とし、研究代表者を含めて多くの試みが行われてきた。しかし和声進行を規則化しようにも和音の切れ目と和音名の同定が一般には困難であり、これはカデンツ認識に支障をきたす。さらに、音楽は言語のように厳格な生成の規範があるわけではなく、局所的な進行の関係を維持すれば比較的自由的な進行が可能である。このよう

な自由的な進行を寛容に捉えるパーサは逆に実装困難である。

2. 研究の目的

まず音楽一般に構成的・階層的な構造があることを主張するならば、その階層的な構造を生み出す生成規則として文脈自由文法(Context-Free Grammar; CFG)を仮定すべきである。CFGは遠隔の依存関係(係り受け)を支配する規則であり、文が係り受け非交差であることとCFGで生成できることは同等である。さらには音楽が言語と出自を一にし、脳内の短期記憶をプッシュダウンスタックとして使うことから音楽認識にCFGを用いる根拠がある。しかしながら音楽が厳格な生成規則によって構成されると考えるのは不可能である。本研究では統計的学習による制約の緩いCFGの発見を考察する。

得られた文法学習モデルは隠れ文脈自由規則モデルとして一般化した枠組みで提示し、認知科学的検証を行う。本研究の独創的な点は以下に集約される。

(1) 文法が生成論的(generative)であるとはチョムスキー(N. Chomsky)の考えであり、有限の生成規則からあらゆる文法的な文が導出可能であることを言う。われわれは音楽にもおいても同様な考えを導入する。CFGによる規則を用いることで数多くの自然言語処理分野で蓄積された知見の借用が可能である。すなわち言語理論を記述するために開発されてきた属性構造による文法HPSGやカテゴリ文法などの文法理論に加えて、パーサの構築、意味記述装置による音楽の調の同定を試みる。

(2) 音楽の進行を司る理論において、隠れマルコフモデルは実世界データ(遷移確率)に基づく規則のリアリゼーションという意味で意義深い。遠隔の依存関係を説明できないという欠点がある。一方で木構造を形成するGTTMも、和声解析ができない限りはこの依存関係が反映されない。本研究ではこれら先行する理論の欠点を補い、依存関係を表現する枠組みを構築する。文生成の確からしさを見るには確率付き文脈自由文法(PCFG)が有用であるが、PCFGだけでは音楽の特有の経過句の自由さを許容することができない。われわれが本研究でめざす生成規則は統計的確からしさを含み、かつシンボル列に局所的な制約を設ける緩い文脈自由規則である。

本研究の構想は後の生成論的音楽理論の構築に向けての準備にあたる。これまで音楽家たちの言う音楽理論というのは数理的考察には程遠い。GTTMを含めいくつかの理論は形式化・抽象化を求めたが、今こそ計算科学で扱うことのできる音楽理論があつてしかるべきである。ここで生成論的音楽理論の完成とは形式文法によるシンタクスの規定と素性構造などによる階層的なセマンティクス表示装置を揃えることである。本研究はこの理論構築の基盤と位置づける。

3. 研究の方法

音楽の文法規則は言語ほど厳密な構文を持たず、経過句は自由である。これには DNA の遺伝子間にあるパディングのように、制約の箇所を限定する規則を構築する必要がある。さらには、限られた数のサンプル入力から規則化を行う必要があるため、ある程度ランダムな確率的挙動を許して規則化を行うメカニズムが必要である。したがって文法規則の音楽への応用には、まず musicXML など表現された多様な記譜音列に対し、緩やかな文脈自由規則を推測するシステムを構築し、抽出された規則をもとにパーサを構築して規則の妥当性を評価することとなる。

音楽の構造認識にとって和音列認識は必須である。隠れマルコフモデルでは、「ソシレ」という和音を聞いたときにそれが八長調のVの和音(V/C)であるかト長調のIの和音(I/G)であるかを決定するため(図2)に、和音間の遷移確率を用いた。

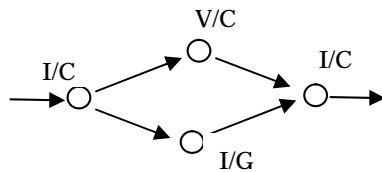


図2 「ソシレ」の和音の調認識

しかしながら和音進行こそは上記の転調を含めてある程度自由な動きを許容し、生成的な規則で記述することが困難である典型的な例となる。したがってその規則を論理的な推論で見出すのは一般には困難である。ここで、われわれは多量の和音列から統計的に規則を発見するため S. Kirby による繰り返し学習モデル(Iterated Learning Model; ILM) による言語伝達のしくみを考察する。ILM では子エージェントは親エージェントから受け取る文字列がどのような文法規則から生成されたものであるかどうかを知ることはできない。しかし子エージェントは文字列を蓄積してそれを一般化することにより、自分で勝手に規則を構成する能力を持つ。子は親のいくつかの発話から、共通部分を変数化する(chunk と呼ばれる)方法により構成的な文法を作り出す過程を示す。この結果、子は親から聞いていない文を発話できるようになる。次に、子エージェントが次世代の親エージェントとなり、新たな子に自分の作成した文法を用いて発話を行う。この過程を何世代も繰り返すことにより、子が発話できる文は増加し、同時に文の生成規則数も構造化されて減らすことができる。

これらの規則を収束させるためには、規則化のアルゴリズムを洗練する必要がある。例えば、戦略のないアドホックな chunk と繰り返し学習では、収束も遅くまた十分に構成的な規則が得られるかどうかは保障されない。

共通部分抽出にも楽句の最長一致の是非や句頭から/句末からの一致戦略など検討課題も多い。本研究では chunk を行う頻度・タイミングを含めてアルゴリズムを洗練し、効率化を試みる。音列・和音列の切れ目推定においては最大エントロピー法が活用できる。またコストを入れるなら Viterbi アルゴリズム/EM アルゴリズムとの組み合わせが考えられる。

また一方で、構文規則を機械的に作るだけならファイル圧縮のための文字列構造化アルゴリズム Sequitur, LZ78, Re-Pair などが存在するが、これらが人間の認知構造を反映しないのは明らかである。本研究では、人間が計算機に対して行う操作列にこのような機械的なアルゴリズムを適用し、得られた規則に認知的妥当性について検討を行う。

4. 研究成果

本研究は次の三つの方向からアプローチを行った。

- GTTMにおける距離の概念の導入と類似性
- ILMにおける言語間距離の評価
- 計算機の操作列からの CFG の発見

これらの着目点について、成果を年度毎に以下のようにまとめる。

(1) 平成 25 年度

音楽にとって文法概念に沿うものとして着目するのが和声進行の規則と拍節構造に基づく木構造である。この観点より 25 年度は以下の研究実績を得た。

Tonal Pitch Space (TPS) の実装: 和声進行の規則を特定するためには、同時に鳴っているいくつかの音をまとめて和音を認識する必要があるが、和音特定は前後の和音との接続性に依存する。TPS は和音間距離を測る理論であるが、この計算機構を実装した。

② ピッチ選好を考慮したタイムスパン木: 拍節構造に基づいて作成した木構造(タイムスパン木)にはピッチ(音高)の概念が欠如しているために、調というコンテキストの中での音の遷移のスムーズさという視点が盛り込まれていない。本年度はタイムスパン木に音高の優先度の概念を持ち込んだ理論を提案し、一部システムに実装した。

③ カデンツ発見システム: 音楽の文法としてもっとも規則性が高いのがカデンツ(和音進行の終止形)である。われわれはこれまでの研究でタイムスパン木の自動化を行ってきたが、カデンツ処理は未了であった。本年度よりこのシステムの実装を開始し、サンプルデータベースにある曲のカデンツ認識システムを構築した。

計算機に対する人間の操作もスタックによる短期記憶を用いた係り受け構造があると仮定できる。この考え方にに基づき、操作列を記憶しておき、Sequitur によって文脈自由文法化する試みを行った。

人工知能学会オーガナイズドセッション: 音楽にも言語と同様に文法構造を見出

せることを主張するためには、この前提として、人間の間のコミュニケーションにおいては木構造を渡すこと、あるいは渡した一次元の文字列から共通した木構造を構築できることが意志疎通であると考え、われわれはこのより大きな目標のために(社)人工知能学会全国大会においてオーガナイズドセッションを企画し、コミュニケーションについて討論する場を設立した。

(2) 平成 26 年度

これまで取り組んできた木構造の構成に関し、距離の概念を改善することにより、木構造がより直截的に生成規則と結びつく試みを行った。

② 文法規則を発見する手段として ILM を用い、言語進化の実験を行った。まず木構造の構成についてはこれまで GTTM に基づくボトムアップなプロセスを主にしていたが、この構成規則は生成的とは言い難い。このため、本年は生成された木がより正しく人間の認知する楽曲の骨格を反映するかどうかを検討するため、前年度に行った構造間の距離の概念を拡張し、ピッチの情報を取り入れた。

③ モーツァルト「キラキラ星の主題による変奏曲」K.265/300e の 12 個の変奏曲の間の距離を計測し直し、MDS (多次元尺度構成法) によって心理距離との差が改善されたことを確かめた。さらに変奏曲間でできた木を組み合わせて人工的な変奏曲を構成し、これがオリジナルの変奏曲とどのように近いかも計測し、心理距離の比較を行った。

繰り返し学習モデルによる世代間の言語間距離の測定を行い、社会ネットワークの形状による情報伝達のモデルに応用した。また文法学習の際に不必要に長い文字列を持った語彙規則が現れるため、覚悟の語尾を短縮することによって文法学習を高速化させることに成功した。

文法を記述する枠組みとして Combinatory Categorical Grammar (CCG) によるアプローチを開始した。

(3) 平成 27 年度

CCG を応用した楽曲パーサの開発: 近代的な和声(すなわちジャズ和声)を含むパーサに対し、先行研究はすべての和声がトニックへの解決を希求するなど無理な設定があった。本研究では和声の機能を拡張してこの理論を発展させた。

② 木構造のサブ構造に基づく類似性: これまで木の類似性はタイムスパンの累積和で定義してきたが、一カ所でも右分岐・左分岐があると距離が大きくなってしまい直観に反していた。この方法を改善し、部分木ごとに距離を積算するしくみを考察した。

③ フリップ操作による楽曲の逆行表現: 楽曲構造を表す木構造においては左右の分岐を入れ替える(flip)操作によって新しい編曲のアルゴリズムを考案した。特に全分岐のフリップはメロディの逆行に相当することを証明した。

確率文脈自由文法(Probabilistic Context-free Grammar; PCFG)を応用した楽曲のタイムスパン木解析: タイムスパン木がほぼ文脈自由文法に沿うことは観察されていたが、その規則の生起確率を計算することにより、解析の精度を向上させることができた。

タイムスパン木におけるカデンツの認識システム: これまでの楽曲のタイムスパン解析においては楽曲の終結部におけるカデンツの保持(cadential retention)が未着手のまま取り残されていた。この研究ではこのカデンツ発見の鍵となるドミナントの和音の認識について、従来のドミナントの和音の解釈を拡張し、カデンツ候補を純粋に増やすこととし、結果として精度を大幅に上げた。

これまでの楽曲解析における文法的手法をまとめ、構造的楽曲解析の著書 2 章分を担当執筆し、出版した。

5. 主な発表論文等

[雑誌論文](計 2 件)

(1) S. Tojo, K. Hirata, M. Hamanaka, Computational Reconstruction of Cognitive Music Theory, New Generation Computing, vol.31, no.2, pp.89-113, 2013(査読有)

(2) S. Tojo, K. Hirata, Distance and Similarity of Time-span Tree, Journal of Information Processing, vol.21, no.2, pp.256-263, 2013(査読有)

[学会発表](計 23 件)

(1) M. Hamanaka, K. Hirata, S. Tojo, Structural Similarity Based on Time-span Subtrees, 5th Mathematics and Computation in Music, London, UK, June 22-25, 2015.

(2) R. Matoba, M. Nakamura, S. Tojo, Simulation of the Emergence of Language Groups Using the Iterated Learning Model on Social Networks, COGNITIVE2015, Nice, France, March 22-27, 2015.

(3) R. Matoba, M. Nakamura, S. Tojo, Application of Loose Symmetry Bias to Multiple Meaning Environment, COGNITIVE2015, Nice, France, March 22-27, 2015.

(4) M. Hamanaka, K. Hirata, S. Tojo, sigma-GTTM III: Learning based Time-span Tree Generator based on PCFG, CMMR2015, Plymouth, UK, June 16-19, 2015.

(5) E. Benjamin Luke, 棟方 渚, 小野 哲雄, 楽曲の表層構造解析に基づくリスナーの生体反応の考察, 第 29 回人工知能学会全国大会, はこだて未来大学, 北海道函館市, 6/6-9, 2015.

(6) 趙 越, 福田 玄明, 植田 一博, 音楽の局所的变化が全体的印象に与える影響の分析, 第 29 回人工知能学会全国大会, はこだて未来大学, 北海道函館市, 6/6-9, 2015.

(7) 小玉 昂史, 東条 敏, モード依存の Tonal Pitch Space, 第 29 回人工知能学会全国大会,

はこだて未来大学, 北海道函館市, 6/6-9, 2015.

(8) 平田 圭二, 東条 敏, 旋律の逆行とタイムスパン木に対する flip 操作, 第 29 回人工知能学会全国大会, はこだて未来大学, 北海道函館市, 6/6-9, 2015.

(9) 松原 正樹, 東条 敏, 平田 圭二, GTTM タイムスパン木における構造レベルの導入と代数表現, 第 29 回人工知能学会全国大会, はこだて未来大学, 北海道函館市, 6/6-9, 2015.

(10) R. Matoba, H. Sudo, M. Nakamura, S. Hagiwara, and S. Tojo, Process Acceleration in the Iterated Learning Model with String Clipping, ICCI2014, Paris, France, December 6-7, 2014.

(11) K. Hirata, S. Tojo, M. Hamanaka Algebraic Mozart by Tree Synthesis, International Computer Music Conference (ICMC/SMC), Athens, Greece, October 22-26, 2014.

(12) M. Matsubara, K. Hirata, S. Tojo, Distance in Pitch Sensitive Time-span Tree, International Computer Music Conference (ICMC/SMC), Athens, Greece, October 22-26, 2014.

(13) M. Hamanaka, K. Hirata, S. Tojo, Musical Structural Analysis Data Based on GTTM, The 15th International Society for Music Information Retrieval (ISMIR), Taipei, Taiwan, October 27-31, 2014.

(14) 外谷 弦太, 橋本 敬, 階層構造生成能力の進化シミュレーション, 日本人間行動進化学会第 7 回大会, 神戸大学, 兵庫県神戸市, 11/29-30, 2014.

(15) 平田 圭二, 東条 敏, パーンスタイン「答えのない質問」再考, 第 28 回人工知能学会全国大会, ひめぎんホール, 愛媛県愛媛市, 5/12-15, 2014.

(16) 松原 正樹, 東条 敏, 平田 圭二, 音楽理論 GTTM に基づく木構造を用いた旋律の認知的類似度の導出, 第 28 回人工知能学会全国大会, ひめぎんホール, 愛媛県愛媛市, 5/12-15, 2014.

(17) 小玉 昂史, 東条 敏, exGTTM を用いたカデンツ部の探索, 第 28 回人工知能学会全国大会, ひめぎんホール, 愛媛県愛媛市, 5/12-15, 2014.

(18) E. Benjamin Luke, 棟方 渚, 小野 哲雄, 和声学に基づく自動作曲システム CMY, 第 28 回人工知能学会全国大会, ひめぎんホール, 愛媛県愛媛市, 5/12-15, 2014.

(19) A. Marsden, K. Hirata, S. Tojo, Towards Computable Procedures for Deriving Tree Structures in Music: Context Dependency in GTTM and Schenkerian Theory, SMC (Sound and Music Computing Conference), Stockholm, Sweden, July 30-August 3, 2013.

(20) K. Hirata, S. Tojo, M. Hamanaka, Cognitive Similarity based on Tree

Distance from the Analysis of K.265/300e, 10th International Symposium on Computer Music Multidisciplinary Research (CMMR), Marseille, France, July 30-August 3, 2013.

(21) M. Hamanaka, K. Hirata, S. Tojo, Time-Span Tree Analyzer for Polyphonic Music, 10th International Symposium on Computer Music Multidisciplinary Research (CMMR), Marseille, France, July 30-August 3, 2013.

(22) R. Hatano and S. Tojo, Detecting Context Free Grammar for GUI Operation, 8th International Conference on Knowledge, Information, and Creative Support System (KICSS), Krakow, Poland, November 7-9, 2013.

(23) 東条 敏, 進化言語学における認知バイアスの有効性, 第 27 回人工知能学会全国大会, 富山国際会議場, 富山県富山市, 6/4-7, 2013.

〔図書〕(計 2 件)

(1) K. Hirata, S. Tojo, and M. Hamanaka, An Algebraic Approach to Time-Span Reduction, David Meredith (Ed), *Computational Music Analysis*, Chapter 10, pp.251-270 (20 pages), Springer (2016).

(2) M. Hamanaka, K. Hirata, and S. Tojo, Implementing Methods for Analysing Music Based on Lerdahl and Jackendoff's Generative Theory of Tonal Music, David Meredith (Ed), *Computational Music Analysis*, Chapter 9, pp.221-249 (29 pages), Springer(2016).

6. 研究組織

(1) 研究代表者

東条 敏 (TOJO, Satoshi)
北陸先端科学技術大学院大学・情報科学研究科・教授
研究者番号: 90272989

(2) 研究分担者

小野 哲雄 (ONO, Tetsuo)
北海道大学・情報科学研究科・教授
研究者番号: 40343389

植田 一博 (UEDA, Kazuhiro)
東京大学・大学院情報学環・教授
研究者番号: 60262101

橋本 敬 (HASHIMOTO, Takashi)
北陸先端科学技術大学院大学・知識科学研究科・教授
研究者番号: 90313709

平田 圭二 (HIRATA, Keiji)
はこだて未来大学・システム情報科学部・教授
研究者番号: 30396121