

人間の言語機能と音楽

東条 敏^{1,a)}

概要: 人間は音楽と言語の両方を持つが、両者は生物学的には同ルーツである。すると、自然言語を扱う手法で音楽を解析することで、効率よくかつ過去に例のないような楽曲解析が可能になる可能性がある。特に文脈自由文法による解析では、楽曲に対して木構造を付与することができる。これにより、音楽をより計算論的・科学的に解析することができるとともに、編曲や検索など現実的な応用が可能である。

キーワード: 楽曲解析, 木構造, 文脈自由文法, 自然言語処理

Human Language Faculty and Music

SATOSHI TOJO^{1,a)}

Abstract: It is a known fact that the origin of music and language is biologically one and the same. If we are to apply various methodology for natural language to music structure, we may realize more efficient and more innovative analyses. The parser by context-free grammar organizes a tree structure, also for music. Therefore, we would be able to consider music in more structural and computational way, and in addition, to implement practical systems such as arrangement or search engine.

Keywords: Music analysis, tree structure, context-free grammar, natural language processing

1. コンピュータには何ができないか

『コンピュータには何ができないか』は米国の哲学者ヒューバート・ドレイファスが1970年代から90年代にかけて三度に渡って版を重ねた人工知能批判の書である。同書は機械による人間性の模倣を否定し、道徳性・愛・創造性は機械の限界を超えているとした。同書によれば、機械に解決可能な問題は選択肢が限られた閉世界にのみあり、機械は問題解決の快感を知ることもないとする。しかしこのような議論は遠くギリシア時代にまで遡る。プラトンは還元主義を盾にとり、機械だって時間と十分な小ささの素子があれば人間のすべての所作・芸術作品の創作も可能であると考えた。時間は隔たるが、中世の哲人デカルトは、逆に人生で起こりうることをすべてに應じるに十分なほど多種多様な装置を機械して備えることは事実上不可能であると考えた。

しかしドレイファス本の後30年を経て、状況は変わりつつある。ほんの30年前までは、コンピュータから導き出せるのは人間が与えた仮定とデータと、そしてコンピュータに内蔵された推論規則によるデータの組み合わせであるとされた。その当時なかったのはウェブである。今大きく変わったのは、一点めには、ウェブの上には膨大な電子データがあり、コンピュータが勝手にそれを収集することができることから何が入力されているのか人間にはわからなくなったことである。そして二点めには、推論規則と呼ばれていたものもプログラミング言語で言うところのif-then-elseではなく、ニューラルネットワークの結合重みが変わることにより、柔軟に入力と出力の関係を変更できるものとなったことである。その結果、コンピュータの内部で行われる計算メカニズムが人間に見えないものとなり、ブラックボックス化されてしまった。

今やコンピュータが人間の愛や創造性を模倣できるかどうかという議論はナンセンスである。デカルトが言うように人間のすべての所作を実現することはたぶん不可能であ

¹ 北陸先端科学技術大学院大学, 石川県能美市旭台 1-1

^{a)} tojo@jaist.ac.jp

ろうが、同時に人間には知り得ない「機械知」というものが実現されることになるだろう。そしてそれにより人間が今まで出会ったことのない、しかし人間にとっては十分に「愛」や「創造性」とわかる概念が機械によって提示されるかも知れない。

さて、音楽は人間固有の知的活動である。コンピュータは、聴覚がないがゆえに、音楽を理解しその良し悪しを独自に評価できるとは考え難い。ところが、十分な数の音楽データの入力と計算機独自の推論機構によって、人間にも十分満足できる音楽が生成される可能性は否定できないのである。

幸いなことに自然言語処理には長い歴史と研究成果の蓄積があり、もともと自然言語を持たないコンピュータにとっても、人間の言語機能を模倣する努力が行われてきた。本稿の目的は、自然言語処理の方法論を音楽にも適用しようというプロパガンダである。これにより、コンピュータが音楽を理解し、あるいは創造（作曲）するプロセスを模倣できるかどうか考察することを企図する。

2. 音楽と自然言語

2.1 音楽の言語への従属

近年英語教育の前倒し化が進み、幼児の英語教育が盛んになった。昨今よく聞くことばは英語のリズムを学ぼうという言い方である。我々が自国語をしゃべるとき、外国語を学ぶとき、その言語らしさを出すためには適切な抑揚とリズムが必要である。そして適切な母音の長さがある。日本語にはモーラという概念があるおかげで五七五と言えば俳句、七七がつけば短歌となり、各音の時間的長さは一律に保たれる。このモーラは英語のシラブル（音節）とは別物である。子音で終わる英単語の最後に母音をつけてしまうといわゆるカタカナ英語になり、英語には聞こえない。ここで気づいて頂きたいのは、抑揚とはピッチの変化であり、ピッチ・リズム・長さというのは音符の概念であるということである。この意味で、音楽は既に自然言語に組み込まれている。

音楽と言語は生物学的には同ルーツであるとされる。咽喉を使って発声し耳を使って聞く点で、生物学的なデバイスは共通である。人間だけが音楽と言語と両方を持つ。動物の発声、特に小鳥の言語を「歌」というのは単なるメタファーであり、あくまで彼らは意志を伝達し合っているのである。言語は意味を持ち、それゆえにコミュニケーションの手段となる。音楽はどの時代・どの文明にも遍く存在するが、それ自身の意味を持つとは言い難く、今となっては日々の生活の何の役に立っているのか明確な説明はない。ダーウィンは人間にとっての音楽は太古の言語の名残りであると考え、男女の間で愛を伝える手段であったと考えた。すると、長い生物学的進化の過程では、人間は音楽を失っている際中なのかも知れない。

生物進化とともに今度は、文明期以降の文化進化を考えよう。音楽 (music) の語源はギリシアの女神ムーサ（英語名 Muse）であり、そこから派生したムーシケー（*μουσική*）がミュージックになった。このムーサは芸術・舞踏・音楽などを司る神であったので、ムーシケーも同様に舞踏・音楽、そして詩歌の詠唱までを含む概念だった。このように歴史的にも音楽という概念は言語（詩歌）と同居した。音楽を発展させたのは、どの文化においても祭礼であり、宗教的儀式である。太鼓のリズム、全員の唱和、そして舞踏である。とりわけ西洋調性音楽（いわゆるクラシック音楽）の発展期においては、音楽は教会の中で神への祈り（言語）を捧げる際に敬虔さを増幅させる装置として機能した。

以上のように考えると、どうしても音楽は言語に対して従属的な地位に置かれる。いずれの時代・いずれの地域においても、音楽は言語で伝えるべき内容をより効果的に提示するための補助手段のようである。現在、ポピュラー音楽が全盛である。多くの場合、歌詞があり、音楽はその伴奏としての地位にある。器楽だけで芸術の高みに達したのは、むしろ例外的に、18~19世紀のクラシック音楽から20世紀のジャズに至るころまでの一時期に限られているように思える。

2.2 自然言語と形式言語

ノーム・チョムスキーは変形生成文法・GB理論を経てミニマリストの理論に至るまで、自然言語の文法理論に貢献した。一方で、その言語の複雑さを形式言語の階層として定義した。例えば、ジュウシマツの発話（単文）は有限状態オートマトンで模倣できることが知られており、それゆえジュウシマツの言語はチョムスキー階層における正規文法 (Regular Grammar; RG) であると考えられる。一方、人間の言語は主語-述語、形容詞-名詞、冠詞-名詞句など多様な側面で係り受けが見られ、かつ文の中に文構造を節として再帰埋め込みができることから、いくつかの例外を除けばほぼ文脈自由文法 (Context-Free Grammar; CFG) であると考えられる。ここで、同文法を定義する文脈自由規則とは、左辺が単一の変数（非終端記号）・右辺が定数（終端記号）と変数の任意の列である。左辺の変数の出現位置がその両側にある記号、すなわち文脈の情報に依存しないことが文脈自由の名前の由来である。自然言語にも文脈自由ではない例外はいくつか知られている。人間の自然言語のクラスを文脈自由文法と文脈依存文法の間でより精密に特定しようという試みも今行われている。

文脈自由文法は非決定性のプッシュダウン・オートマトンによって模倣できる。文脈自由規則は Greibach 標準形（右辺の先頭一文字を定項、残りを非終端記号列とする）にすることができる。この右辺の定項一文字を読み込んだ時に、規則左辺の非終端記号がプッシュダウン・スタックのトップにあれば、それがポップされ、右辺の残りの非終端

記号列がプッシュされると考えればよい。ここでスタックに積む記号を開き括弧とし、係り受けのものの語を開き括弧、係り受けの先の語を閉じ括弧と考えれば、プッシュ・ポップの操作を括弧の開閉に対応づけることができ、結果として係り受けが交差しない仕組みを作ることができる。

2.3 音楽の木構造解析

文脈自由規則は左辺の非終端記号が右辺に展開される構図を上下に並べて、全体として木構造を作る。したがって自然言語はその構文を木構造で表示することができる。さて、音楽と言語の出自が同じなら、音楽にも同様の木構造が存在するだろうか。

文法理論を用いて和声進行規則を解析しようという試みはテリー・ウィノグラード (1968) を嚆矢とし、指揮者・作曲家として有名なレナード・バーンスタインは楽曲の構造をチョムスキーの生成文法になぞらえてハーバード大学で講義を行った (1973)。近年ではカテゴリー文法を用いたジャズ和声の解析の試みなども見られる。これらはいずれも和声の進行規則を文脈自由規則と捉え、木構造を生成するモデルである。

一方で、チョムスキー的な階層構造を意識しながら、文脈自由規則を用いずに階層的な木構造を生成する理論が Generative Theory of Tonal Music (GTTM) である。この理論においては、曲内の各音は楽曲を構成する上で重要度に差があり、局所的な楽句の両端と強拍の音はより重要であるとする。こうして隣接する音どうしの重要度を比較すると、より重要な音が隣の音 (あるいは隣の部分木) を支配するという関係を生む。すると、ボトムアップな勝ち抜き戦によって木構造が構築できる。そして、木の階層のあるレベルより上を見ると、20世紀初頭ハインリッヒ・シェンカー (1868–1935) が提案したような、楽曲の基本構造が抽出できると考えられる。我々はこの理論をコンピュータ上に実装し、楽譜情報を入力すると構文木を出力する解析器を構築した [1]。この GTTM 解析器によっては楽曲検索、自動編曲、類似度計算、二曲のモーフィングなどさまざまな応用が可能である。

さて、GTTM においても文脈自由規則による係り受けの機構を組み込む局面がある。まず、GTTM においては全体構造を「和声的に安定」させるために局所構造の再編を行う必要がある。例えば局所的に見てより重要な音でも、全体構造からみて従属的な地位に再配置されることがある。これは先行するピッチイベントの係り受けの「受け」であり、これにより、すぐ隣接しない音からの (遠隔の) 依存を起こすのである。このことは、人間の言語理解における直観によく合致する。我々は言語を聞くと、主語があれば将来述語が来るのを期待する。英語で他動詞が来れば目的語を期待する。冠詞が来れば名詞を期待する。同様に、音楽においてはトニックから開始されたと思えば、ドミナ

ントで緊張を感じ、トニックによる終結を期待する。このような「期待」は文脈自由規則が生む部分木に依拠する。

近年の新しい計算機科学のパラダイムにおいては、文脈自由規則とは人間が手で書いて固定したものに拘る必要がない。最初に書いたように if-then-else の構造は柔軟に変更可能なものとみなすべきである。従って文脈自由文法も確率付き文脈自由文法 (Probabilistic Context-Free Grammar; PCFG) で扱うほうが妥当であろう。また、確率的な動きをサポートするのであれば、表層的に表れているシンボル列 (この場合は音符列・和音列) を n -グラムとして扱うだけでなく、中に意味的にふるまう隠れ状態 (この場合は和音機能など) があると仮定する隠れマルコフモデル (Hidden Markov Model; HMM) も有用であろう。すると、楽曲の解析は近年自然言語処理で行われている成果が大いに役立つのではないだろうか。

3. コンピュータの創造性

今日、人工知能は言語の意味がわからずとも、ある精度で機械翻訳ができるようになりつつある。コンピュータは言語の意味はこの世界の何物にもグラウンディングすることなしに、このプロセスを実現しているのである。それならばコンピュータは味覚がなくともウェブからアルコール飲料のデータとそれに対する評価を集めてきて、人間の満足するカクテルを作る可能性があるかも知れない。同様に、コンピュータは、聴覚に基づく審美眼を持たなくても、何が人間を満足させているのかという評判だけを集めてきて、人間の感性に合致する音楽を作曲することができるだろうか。さらには、我々の美学に応える演奏を行うことが可能だろうか。

謝辞 本稿の研究内容は JSPS 科学研究費 基盤 (A) 16H01744 の支援を得たものである。

参考文献

- [1] Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo: "Implementing A Generating Theory of Tonal Music", Journal of New Music Research (JNMR), Vol.35, No.4, pp.249-277, 2007.
- [2] 東条 敏, 平田圭二: 音楽・数学・言語。近代科学社, 2017.