

メロディの調を認定する過程の計算モデル

吉野 巖

北海道大学大学院文学研究科

yoshino@hubs.hokudai.ac.jp

阿部 純一

北海道大学文学部

abe@hubs.hokudai.ac.jp

メロディの調が認定される認知過程の計算モデルについて述べる。メロディを入力すると、このモデルは、知識内の調性的スキーマ（調性構造に関する知識）を援用しながら調決定を行う。心理学の文献から、平均的な西洋音楽の聞き手の調性的スキーマは、長音階と短音階の主音を中心に、主和音構成音、その他の全音階構成音、などが階層的に構造化されたものである、ということが示唆されている。また、聞き手が入力された音列を実際にどのような調のメロディとして認定するか、を調べる心理学的実験を行ったところ、提示された音列を、調性的スキーマのより上位の階層にあてはまるような調のメロディとして解釈する傾向のあることが確認できた。本モデルは、このような認知的特徴を計算機上に実現したもので、上記の実験結果を良く予測した。

Computational model of the process of the "key" determination.

Iwao Yoshino and Jun-ichi Abe

Department of Behavioral Science
Faculty of Letters
Hokkaido University

This paper describes a computational model which interprets a key of given melody. The interpretation is guided by a "tonal-schema model" provided in its knowledge-base. Some psychology literatures suggest that western-music listeners' tonal schema is constructed by the hierarchical representations of tonic center (so-called tonic-tone), elements of tonic triads, and other diatonic scale tones. A psychological experiment was done to obtain the data of listeners' real performance of key interpretation. The results indicated that listeners tend to interpret a key of a given tone-sequence as to assimilate to their internal tonal schema which have the properties mentioned above. This model implemented an internal "tonal-schema model" which have such cognitive properties, and could predict the experimental results well.

1. はじめに

メロディを聞く時に我々が感じる“まとまり”や“メロディらしさ”、また“記憶しやすさ”などは、西洋音楽でいう“調”と関係のあることがわかっている。つまり、普通の人でも、調性構造に関する知識（調性的スキーマ: tonal schema）を心内にもっており、それに基づいて無意識的に調を認識していると考えられる。

調性的スキーマに基づき、メロディを構成する個々の音高を心的に体制化することは、“調性的体制化 (tonal organization)”と呼ばれる。我々は、この過程の処理の結果として得られるものが、いわゆる“調”ではないかと考えている。その理由は、調性的体制化の過程とは、本質的に、調性的スキーマに基づいて参照点的役割を果たす“中心音 (tonal center)”と他の音とを関係づける過程であり、中心音が定めれば調が定まったことになるからである（阿部、1987b）。

調性的体制化の過程については、まだ、十分に精密でかつ心理学的に妥当性の高い説明理論が提案されていない。本研究では、“調”を認定する計算モデルの構築を目指す。その作業を通じて、人間の調性的体制化の過程についての理解を深めようというわけである。

2. 西洋音階の階層構造

西洋音楽には長音階と短音階の2種類の音階がある。それぞれの音階を半音単位の音程で表すと、長音階

{2, 2, 1, 2, 2, 1}、短音階 {2, 1, 2, 2, 1, 3, 1} ということになる。どちらの音階も7音階で極めて非対称的な構造特性をもっているため、位相を1オクターブずらさない限り構成音が一致することはない。

ある音高の機能（つまりは主音との関係）はその調によって異なってくる。その機能は図1中に示したような階名として表すことができる。例えば長音階はどの調でも中心音がdoで、それ以外は re、mi、fa、sol、la、si という機能をもった音から構成されると見なすことができる。音高C3はC-majorではdoという機能をもつが、G-majorではfaという機能をもつ。一方、短調は中心音がlaで、それ以外は si、do、re、mi、fa、sol# という音からなる。前述のC3は同じ音高であっても、C-minorではlaという主音としての機能をもつことになる。

伝統的な音楽学では、西洋の全音階 (diatonic scale) の構成音の相対的な重要度に階層性があるとされている。この階層の最も上位に位置する音が主音 (tonic) であり、最も下位に位置するのが5つの非全音階音（長調の場合は、do#, re#, fa#, sol#, la#, 短調の場合は、la#, do#, re#, fa#, sol）であることは、経験的にも明かである。全音階音の中では、主音の次に、主和音（I度音）の構成音である3度音 (mediant、長調のmi・短調のdo) と5度音 (dominant、長調のsol・短調のmi) や4度音 (subdominant、長調のfa・短調のre) が優越的機能を持つと言われており、それ以外の全音階音（長調: re、la、si、短調: si、fa、sol#）は相対的な優越性が低いとされる。なお、主和音構成音以外の4つの全音階音は、主和音の次に重要な和音で

The image shows two musical staves. The top staff is for C-major, with notes do, re, mi, fa, sol, la, si, do. The bottom staff is for C-minor, with notes la, si, do, re, mi, fa, sol#, la. Above the C-major staff, three triads are shown: I (do, mi, sol), IV (fa, la, do), and V (sol, si, do). Above the C-minor staff, three triads are shown: I (la, do, mi), IV (re, fa, la), and V (sol#, si, do). The notes do, mi, sol in the C-major triads and la, do, mi in the C-minor triads are circled.

図1. C-majorとC-minorの音階構成音と主要3和音（丸で囲った音は主和音を構成する音である）。

あるIV度音とV度音の構成音であるので、全音階音は全て主要3和音（I、IV、V）の構成音であることになる。

このような、経験的にとらえられてきた全音階音の階層性は、心理学的実験の結果からも支持されており（Krumhansl & Kessler, 1982, など）、調性的スキーマがそうした階層性をもっていることが推測される。また、当然のことながら調性的体制の過程もそうしたスキーマの性質を反映したものとなると考えられる。実際、星野・阿部（1984）やAbe and Hoshino（1990）の研究結果は、音列が与えられたときの聞き手の調解釈がそのような性質を示唆するものであった。

3. 調性認定に関する心理学的実験

聞き手が、入力される音列の調をどのように解釈・認定するか、その過程を調べるための実験を行った。この実験で得られた結果は、後にモデルのシミュレーション結果と比較される。

方法

被験者は絶対音感をもつ音楽経験者4名であった。そのような被験者を用いた理由は、“調”という概念が良く理解されており、かつ、音名によってその調を答えてもらうことができる、ということによる。刺激音列としては、バッハの平均律ピアノ曲集の中からの24種のフーガ主題と、新たに作成した10種の無調的音列を用意した。ただし、今回は、音列のリズムの特徴による効果を分析の対象外とすることにし、音列の全ての音符の長さは一律にした。

被験者には、メロディの冒頭から音が1音づつ増える度に、その音列に対する調の認定を行うよう教示した。この手続きは、調認定の漸進的（incremental）な側面を観察するために採用された（図2参照）。出だしの1音のみを与えた場合、被験者はその音を主音とする調に認定することが予備実験の結果から確認されている。そのため、被験者には冒頭の2音に対する調認定から始めてもらい、その後、音が1音づつ加わる度に新たな調認定をしてもらうこととした。なお、認定する調は複

数でもよいし、無調という判断をすることも可能であるととした。



図2. 刺激音列（フーガ17）の提示方法。(1)→(2)→(3)→の順で各音列毎に調認定を行う。

結果と考察

予想されたように、被験者の調認定はその多くが長調または短調の全音階の構造と整合した（全反応中、89.1%）。

被験者は出だしの2、3音を聞いただけでも“調”を認定できた。そこで、2音刺激（図2の(1)パターン）について分析した結果、その2音の音程が3度あるいは5度（または転回音程である4度）音程の時に、主和音を構成する音程関係として解釈する傾向が認められた。この傾向が最も強かったのが完全5度音程の時で、これは10通りの解釈ができるにも関わらず、ほとんどの場合（95.6%）は長調の“do-sol”または短調の“la-mi”として解釈された。実験後、被験者に調認定のストラテジーを尋ねたところ、全音階の主和音を構成する音高、つまり長調のdo、mi、solか短調のla、do、miというような音高を手がかりにした、という内観報告が得られた。上記の完全5度の例はこのことを裏付けているといえるであろう。

被験者は音列の進行に従って、調の解釈を変えることがしばしばあった。解釈を変える原因としては以下のようなことが考えられる。大きく考えられるのは全

音階的枠組の制約によるもので、この例としては図3の被験者S3の解釈を挙げることができる。S3は8音目を聞いた段階で、F-majorとして解釈したが、9音目の段階では調をBb-majorに変えている。これは、F-majorで解釈した場合、9音目Eb3が全音階音として解釈できなくなってしまうことによると思われる。

さらに考えられる原因は、主和音構成音の枠組によるもので、その例としては、同じ図3の被験者S1の解釈を挙げることができる。S1は、5音目の段階でF-majorという調を認定した。6音目D3はF-majorでは“1a”として解釈できるにも関わらず、彼は6音目の段階で調をBb-majorに変えている。これは、Bb-majorで解釈した方が音列の多くを主和音構成音（図3中の2重線で示した音）として解釈できることによると考えられる。なお、被験者S3は6音目でもF-majorを認定しているので、この原因はさほど強いものではないかもしれない。

被験者は、1つの試行につき、ほぼ1つの調を認定していたが、冒頭の2、3音の段階では複数の調を認定する場合が多少見られ（平均1.24個）、音列の進行につれてその割合が減少していく（7音目以降、平均1.03個）、という傾向がみられた。

被験者間の調認定は、出だしの3、4音目の段階ではあまり一致しないが、6、7音目の段階になると、ほとんどの場合一致するようになる。表1の例は、冒頭から被験者間の調認定が良く一致している例である。この音列は冒頭の完全5度が♯を主音とする調を強く導き、その後反証がないために、このような被験者間で一致する、しかも安定した解釈が得られたのであろう。

しかし、音列の後半になっても調の認定が被験者間で一致せず、2、3の解釈に分かれる場合も見られた。そうした例を表2に示す。この例では、4人の被験者の調解釈がはっきりと2通りに分かれている。

このようにみえてみると、調とは、メロディを聞いて徐々に聞き手の心内に形成されていくものである。そして、その過程で認定される調は変化し得るものである。また、聞き手によって認定される調が異なったり、複数の調が感じられることもある、ということがわかる。

4. 調認定モデルの構築

我々は、上記のような心理学的実験から得られた知見をもとに、調を認定するモデルを構築している（阿部 1987a; 吉野・阿部 1992; 吉野・阿部 1993, も参照されたい）。モデル構築にあたっての最も基本的な仮定は、人間のメロディ処理の過程を、聴覚的処理の過程と、より高次の認知的処理の過程とに分離できるとしたことである。つまり、聞き手は、低次の感覚的過程において、連続的な音響刺激を、離散的な、オクターブ内12個の音高カテゴリーに符号化する。次に、高次の認知的処理過程では、この12個の音高カテゴリーに対して調性的体制化が行われ、なんらかの“調”が認知される。というように仮定する。本モデルでは、入力音がすでに感覚的な処理を経た後の過程、つまりは高次の認知的処理の過程、をモデル化したものと仮定する。

次の仮定は、この高次の認知的処理の過程がスキーマ依存的な性格をもっており、個人差は、基本的には各自がもつスキーマの違いによってもたらされる、としたことである。調性的スキーマは、個々人が体験してきた音楽の調性構造を内化したものであるといえる。本モデルは西洋音楽の聞き手を想定したため、そのモデルが備える調性的スキーマは、西洋音楽の調性構造を反映したものとなっている。

前述の心理学的実験から得られた調性に関する知見は以下の通りであった。

1. 聞き手は全音階的な枠組にあてはまるように調を解釈する傾向がある。
2. さらには、音列ができるだけ全音階の主和音の構成音を中心とした構造となるように調を解釈する傾向がある。

このような傾向を説明できるように、調性的スキーマをモデル化した。その概略は図6に示されている。

我々の仮定した調性的スキーマは、主音を頂点とする主要3和音に基づく和音フレームからなる階層構造をもつ。2で述べたとおり、全音階の音階音7音は、主和音、IV度和音、V度和音の主要3和音の構成音として表現できる。これら3つの和音フレームは、全体で全音階的枠組として作用し、全音階性の高い調解釈を選び出

ず動きをする。

さらに、主和音フレームが他の2つの和音フレームに優越し、主和音構成音が参照点的役割を果たすような解釈が優先されることになる。つまり、それぞれの和音フレームは、独自の構成音枠に基づいた音要素をすくいとるわけであるが、そのときに他の2つの和音フレームよりも主和音フレームの枠内に音列音が落ちるように解釈するのである。

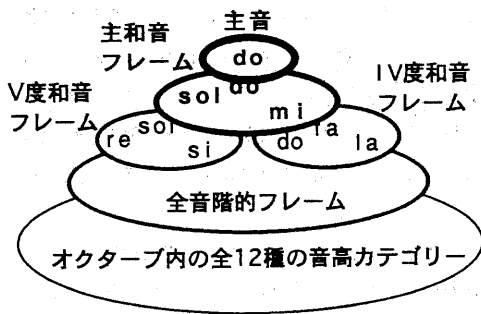


図6. 調性的スキーマ（長調の場合）のモデルにおける階層性。

それぞれの和音フレームは、それ自体の内部にも優越性を持っている。調性的体制化においては、音程が重要な手がかりとなるが、和音の構成音から作られる5度（転回形の4度）と3度の音程は最も重要な音程である。そこで、まず音程関係を考慮にいれ、根音-5度音音程（完全5度）を最も重要な音程とし、その次に根音-3度音音程（3度）、3度音-5度音音程（3度）という順の階層性を持たせた。そして、これら3種の音程に関して、階層の上位のものから順に、3、2、1点の重み値を与えている。また、個々の音高にも和音の根音（root）、5度音、3度音という順の階層性をもたせた。具体的には、根音に3点、5度音に2点、3度音に1点の重み値を与えている。こうした和音フレーム内の優越性に関する概略を図7に示す。

本モデルは、上記のように仮定された調性的スキーマの様々なレベルで、各フレームに対する適合性を評

価する処理が行われ、結果として各調の解釈の可能性が評価値として出力される。具体的にはまず、3種の和音フレームに対する適合度、さらに全音階フレームに対する適合度が計算され、非全音階音を含むような調解釈にはマイナスの重み付けがなされる。こうした処理の結果、24の調全てに対して、その調として解釈できる可能性を表す数値（評価値）が得られる。

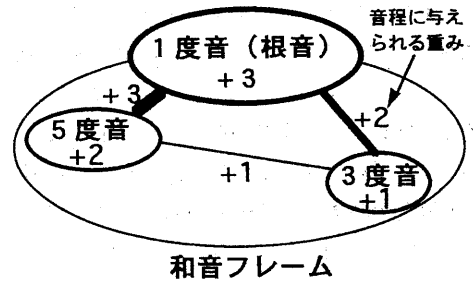


図7. 和音フレーム内の階層性。根音、5度音、3度音の順となっている。3種の音程と3種の音高（和音上の位置を表す“度数”で示してある）は、図中の重み値に反映されるような階層性をもつ。

本モデルは、まず上記のような条件の厳しいフィルタを使用して調を認定しようとする。しかし、入力される音列によっては、どのように解釈しても非全音階音が含まれる場合もあり、その場合は適当な調が見つからないことになる。こうした場合モデルは、フィルタの条件を緩やかにして、非全音階音が1音だけ含まれるように解釈できる調を探す。それでも適当な調が見つからないときは、条件を1段階ごとに緩くしながら探索を続けていく。

さて、ここでモデルのシミュレーション例を見ていくことにする。図8の例では、モデルはまず1st-tone（第1音）の段階で、その音を主音とする調（Bb-major, Bb-minor）を選ぶ。ただし、その評価値は低い。第2音段階で最も可能性が高く見積もられたのはG-minorである。Bb3-G3の音程が短3度ということで、モデルはこの音列が短調の主和音フレームに最も良く適合する（do-laとして）と解釈したわけである。Eb-majorに対する

6. 引用文献

- 阿部純一 (1987a). メロディの知覚と予測の過程: 終止音導出行為のシミュレーション. 心理学研究 58, 275-281.
- 阿部純一 (1987b). 旋律はいかに処理されるか. 波多野誼余夫編, 音楽と認知. 東京大学出版会.
- Abe, J., & Hoshino, E. (1990). Schema-driven properties in melody cognition: Experiments on final-tone extrapolation by music experts. Psychomusicology, 9(2), 161-172.
- 星野悦子・阿部純一 (1984). メロディ認知における“調性感”と終止音導出. 心理学研究 54, 344-350.
- Krumhansl, C.L., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. Psychological review, 89, 334-368.
- 吉野巖・阿部純一 (1992). メロディ認知における調性的体制化の過程—モデルと実験—. 日本認知科学会第9回大会発表論文集, 88-89.
- 吉野巖・阿部純一 (1993). 調(key)決定過程の認知的モデル化: 確率論的モデル, コネクショニストモデル, 記号的モデルの比較と今後の方向性. 日本認知科学会第10回大会発表論文集, 118-119.