

NAIST-IS-DD761027

博士論文

和音進行における音楽的期待の計算論的研究

森本 智志

2016 年 9 月 16 日

奈良先端科学技術大学院大学
情報科学研究科

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士 (理学) 授与の要件として提出した博士論文である。

森本 智志

審査委員：

池田 和司 教授	(主査)
中村 哲 教授	(副査)
吉本 潤一郎 准教授	(副査)
伊藤 貴之 教授	(お茶の水女子大学)

和音進行における音楽的期待の計算論的研究*

森本 智志

内容梗概

音楽聴取時に次時刻の音に対して生じる音楽的期待は、ヒトの自発的な予測に起因する知覚現象のひとつとされる。現代社会において最も普及している調性音楽と呼ばれる楽曲群は、この現象を積極的に利用していることが知られており、音楽知覚研究における重要なトピックとされてきた。これまで、和音進行に対して生じる期待が、先行する連鎖和音の文脈に含まれる調性に影響されることが示唆されてきた。しかし、従来研究では音楽理論と呼ばれる調性音楽の作曲作法を前提とする分析が行われてきたため、調性が明に判定できない連鎖和音については詳細に議論されてこなかった。また音楽理論は経験則の集積に過ぎないため、調性と期待の計算論的な関係が十分に検証されているとは言えない問題があった。本研究では、心理実験のデータに基づいて内的な計算過程をモデル化し検証することで、音楽理論に依存しない分析を実現し、和音進行に対する期待が生じるメカニズムについて調べた。まず、和音進行に対する主観的整合性を問う心理実験から、和音進行に対する期待の持ついくつかの性質を明らかにした。続いて得られた性質に基づいて計算過程をモデル化し、候補のモデル間でどの程度実験データを説明できるかを比較した。その結果、聴取した和音進行から音楽理論上の調に対応する隠れ変数を動的に推定し、推定した調に基づいて和音進行の期待を生じるモデルが、参加者の心理評価をよりよく説明できた。この結果は従来の実験心理学や神経科学研究で得られてきた知見と矛盾なく、音楽的期待が脳の機能的な特性によって生じている可能性を支持するものであった。

*奈良先端科学技術大学院大学 情報科学研究科 博士論文, NAIST-IS-DD761027, 2016 年 9 月 16 日。

キーワード

音楽知覚, 計算モデル, 音楽的期待, ベイズモデル, 音楽理論

A Computational Study on Harmonic Expectancy*

Satoshi Morimoto

Abstract

When one listens to music, expectancy for a coming musical event is formed spontaneously based on a preceding context. Previous studies in experimental psychology have shown that harmonic expectancy is consistent with Western music theory, which provides practical guides to compose and analyze tonal music. Since the music theory does not cover the entire perceptual pattern of expectancy, including short chord sequences and atonal combinations, analyses based on the music theory are not enough to understand the underlying computational processes involved in harmonic expectancy. The aim of this study is to examine the context effects on listeners' chord processing from a computational perspective. At first, several properties of harmonic expectancy were revealed from a behavioral experiment, in which participants listened to short chord sequences and judged the subjective relatedness of the last chords to the preceding ones. Then, the computational process governing the harmonic expectancy was modeled based on the obtained properties. The listeners' judgments were best approximated by a model which holds an internal hidden variable which corresponds to key, and it is updated sequentially by observations. The results suggested that, when listening to chord sequences, internally-constructed tonal assumptions dominate harmonic expectancy.

*Doctoral Dissertation, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD761027, September 16, 2016.

Keywords:

music perception, computational model, music expectancy, Bayes model, music theory

目 次

1	序論	1
2	研究背景	3
2.1	和音と調性	3
2.2	和音進行における音楽的期待の心理学的研究	6
2.3	楽曲構造の情報学的研究	9
2.4	従来研究で残された課題	10
2.5	計算論的アプローチ	12
2.6	本研究の指針	13
3	短い連鎖和音における和音進行への音楽的期待	16
3.1	目的	16
3.2	手法	17
3.2.1	心理実験	17
3.2.2	分析	19
3.3	結果	20
3.3.1	統計的分析	20
3.3.2	多変量解析	26
3.4	考察	29
4	音楽的期待の計算論的分析	31
4.1	目的	31
4.2	方法	32
4.2.1	音楽的期待のモデリング	32
4.2.2	モデルの検証	36
4.3	結果	41
4.3.1	パラメータ推定の精度	41
4.3.2	個人ごとの分析	41
4.3.3	交差検定によるモデルの汎化性能比較	47

4.4 追加分析による検証	52
4.4.1 二値強制選択課題と見なした分析	52
4.4.2 キープロファイルの導出	52
4.5 考察	57
5 総合考察とまとめ	63
謝辞	66
参考文献	68
付録	78
A. 業績リスト	78
B. 補足データ	78
B.1 参加者の個人差について	78
B.2 参加者個人内の DOR のばらつきについて	81
B.3 音楽的期待と音程の基準の関係	83

目 次

1	平均律における 12 種類の長三和音の譜例と名称	4
2	音楽理論を前提とした連鎖和音刺激の譜例	7
3	実験のフロー	17
4	実験に用いた和音刺激の例	18
5	各連鎖和音に対する全参加者の平均 DOR	21
6	各連鎖和音に対する DOR の参加者間の標準偏差	22
7	2 連鎖条件における DOR の分布	23
8	3 連鎖条件における DOR の分布	24
9	3 連鎖条件と 2 連鎖条件の DOR の差分	25
10	多次元尺度構成法による, DOR のパターン分析	27
11	階層クラスタリングによる DOR のパターン分析	28
12	個人ごとのモデル評価のフロー	38
13	交差検定によるモデル評価のフロー	39
14	モデル推定精度検証のための数値実験のフロー	40
15	パラメータの推定誤差	42
16	個人ごとに推定した場合のモデル間の性能比較	43
17	個人分析における BU モデルの推定パラメータ	45
18	個人分析における BS モデルの推定パラメータ	46
19	10 分割交差検定によるモデル間の性能比較	48
20	交差検定により推定されたパラメータ	50
21	BU モデルによる DOR 予測値と参加者間平均 DOR の散布図 . . .	51
22	二値強制選択課題と見なした場合のモデル間性能比較	54
23	二値強制選択課題と見なした分析により推定されたパラメータ . .	55
24	キープロファイルの比較	56
25	期待行列と音楽理論の比較	58
26	個人ごとに推定した期待行列の平均と標準偏差	61
27	DOR 及び期待行列における実験参加者間のばらつき	79
28	BU モデルの期待行列への音楽経験の影響	80

29	2連鎖条件における実験参加者ごとのDORのばらつき	82
30	BUモデルにおける期待と音程の基準の関係	83

表 目 次

1	音楽理論上続きやすい和音進行	5
2	個人分析の試験データに対する相互エントロピーの統計的検定 . .	44
3	交差検定の試験データサブセットに対する相互エントロピーの統計的検定	49
4	期待行列の上位及び下位の和音進行	59

第1章 序論

ヒトは過去の観測に基づいて、次時刻に生じる事象を予測することができる。このとき、ヒトを取り囲む環境は定常ではなく時々刻々と変動するため、予測には環境のダイナミクスも考慮した複雑な情報処理が必要とされる。こうした一連の処理は多くの場合無意識下において自発的に行われており、我々の日常生活を支える重要な高次脳機能のひとつであると言える。

予測を積極的に利用した芸術の一つに、音楽が挙げられる。音楽は様々な音要素の時間的な展開によって主観的な美的観念を生じさせる。このとき、先行する音の文脈に基づいた期待が内的に形成され [1, 2, 3]、さらに協和や非協和といった音響的な特徴が結びつくことで、感情を誘発したり雰囲気演出したりすることができる [4]。今日我々が日常的に接している音楽の多くは、西洋古典音楽に端を発する調性音楽という枠組に分類される。調性音楽は調と呼ばれる知覚的中心を音の文脈上に保持する一方、音要素の連結について内的な期待を積極的に利用する特徴を持つ。こうした調性音楽が文化的な背景の違いに関わらず広く世界中で受け入れられていることは、ヒトの脳がもつ知覚基盤の特性と調性音楽のもつ特徴の間に何らかの関連がある可能性がある。

音楽聴取時に、ヒトの予測的な知覚機能によって生じる後続音に対する期待を、音楽的期待と呼ぶ。音楽的期待は旋律、和声、リズムといった音楽を構成する様々な要素・単位において生じる。20世紀半ばに行動実験とその情報論的な分析が提唱されて以降 [5]、行動心理学において音楽的期待の様々な側面に関する研究が進められてきた。近年は、ヒトの非侵襲脳機能計測も含めた統合的な議論が盛んである [4]。これらの研究は、調性音楽の経験的な作曲上の指針である西洋音楽理論（以下、音楽理論）と、心理評価や脳活動との間に一定の整合性があるという結論を得てきた。しかし音楽理論による解釈は、音楽という複雑な物理刺激を扱う上で数少ない指標となる半面、それ自体を前提とする実験条件の設定や分析に対し、論理面での問題を指摘する批判があった。また、音楽理論が言及していない刺激に対する知覚の研究は困難であるため、打開策となる新たな手法の導入が必要であった。

一方で音楽を情報論的側面から分析する研究が発展し、調性音楽の構造の統計

的な特徴と心理実験で得られてきた音楽的期待に関わる行動選択の特徴との類似が指摘されてきた [6]. このことは調性音楽が音楽理論に従うことと表裏の関係にあると同時に、音楽的期待の統計的な性質を近似できるモデルを設計することが音楽理論の仮定に代わる手法に繋がる可能性を示しているとも言える. 認知や知覚現象に対し、同等の機能を持つ計算モデルを設計することで内的なメカニズムの解明を目指す方法論は、計算論的アプローチと呼ばれる. 先行研究でも計算モデルを用いた行動分析の報告はあるが、音楽理論から導かれる仮定を下地としないモデルによる研究はまだない.

本研究では、和音の連鎖に対して生じる音楽的期待に焦点を当て、その生起メカニズムについて論じる. 特に上述の理由から避けられてきた、短い連鎖和音における期待に注目した. このとき、心理実験のデータから内的な計算過程を直接モデル化することで、音楽理論を前提としない分析手法を実現した. 第2章では、音楽理論に関する予備知識と、心理学及び工学の分野における先行研究を紹介する. その上で問題点を提起し、解決策として提案する行動分析に基づいた計算論的アプローチについて解説する. 第3章では心理実験について説明し、従来研究で用いられてきた手法による分析結果を示して、計算モデルが満たすべき条件について考察を行う. 第4章では前章の結果を参考として計算モデルを設計し、改めてモデルに基づいた行動分析を行う. さらに、推定されたモデルと先行研究との関連性についても検証を行う. 最後に第5章において、結果を総括し今後の課題について述べる.

第2章 研究背景

2.1 和音と調性

本節では、音楽理論に関連するいくつかの用語及び概念について説明する。

ヒトは音すなわち空気振動の周波数に応じて、その高さを知覚することができる。その高さを音高 (pitch) と呼ぶ。また、二種類の音の音高の差を音程 (interval) と呼び、特に 1 : 2 の周波数関係をオクターブ関係という。現代の我々が普段耳にする音楽の多くは伝統的に 440 Hz の音を基準の音高として、オクターブ内を均一に 12 等分した 12 平均律と呼ばれる音を要素として構成される。オクターブの偶数倍関係にある音は同一の種類のように知覚される特徴があり、12 平均律の楽曲では 12 種類のピッチクラス (pitch class) が定義される。隣り合うピッチクラスの音程は半音、半音 2 つ分は全音と呼ばれる。

多くの楽曲は複数の音が様々な形で時間的に展開される。同時に鳴る複数の音で構成される複合音を、和音と呼ぶ。和音は基準となる根音と、その他の構成音との音程関係に基づいて規則的に命名される。特に調性音楽では 3 つのピッチクラスからなる三和音 (triad) が重要視され、根音及び根音から 4 半音と 7 半音上のピッチクラスで構成される和音を長三和音 (major triad)、根音及び根音から 3 半音と 7 半音上のピッチクラスで構成される和音を短三和音 (minor triad) と呼ぶ。例えば、C、E、G のピッチクラスからなる長三和音は、C メジャーコード (C major triad) と表記される。図 1 に根音を最下音とする 12 種類の長三和音の譜例と、それぞれの呼称例 (図 1(a)) を示す。

調性とは、主音と呼ばれる知覚的に中心となる音とその他の音との主従関係からなる体系を指す。主音を根音とする和音を主和音と呼ぶ。調性音楽において、主和音は最も安定した知覚的印象を与えるように位置づけられる。その他の和音が楽曲中で持つ機能的な意味合いは、主和音との相対的な音程関係で決まるとされる [7]。例えば、主音から 7 半音上のピッチクラスを根音に持つ三和音は属和音 (dominant) と呼ばれ、楽曲中で緊張感を知覚させる性質を持つ。調性音楽は調性を明確に知覚させるため、音階と呼ばれる特定の 7 つのピッチクラスを多用する傾向にある。そのため、和音の名称も音階に基づいた表記で記述されることが多

a) C C#
D_b D D#
E_b E F F#
G_b G G# A A#
B_b B

b) I ^bII II [#]II
_{III} III IV [#]IV
_{bV} V [#]V
_{bVI} VI [#]VI
_{bVII} VII

c) 1度 短2度 長2度 短3度 長3度 4度 増4度
減5度 5度 短6度 長6度 短7度 長7度

図 1 平均律における 12 種類の長三和音の譜例と名称

a) 音高に基づく和音名, b) C を主和音としたときの音階に基づく和音名, c) C を基準としたときの度数を表す。

い (図 1(b))。音階上の各音を根音とする和音はギリシャ数字で表記され、それ以外は臨時記号を付けて表される。例えば C を主音とした場合、C メジャーコードは I の和音 (主和音)、F メジャーコードは IV の和音となる一方、F を主音とした場合はそれぞれ、V の和音、I の和音となる。このとき、前者は C メジャーキー、後者は F メジャーキーというように主和音のピッチクラスに応じて調が命名される。また、音程を指す場合にも、慣習的に音度と呼ばれる主音からの音階上における距離に応じた呼称が用いられる (図 1(c))¹。

ある時刻の和音から次時刻の和音への遷移を和音進行と呼び、前後の和音の構成音同士の接続を和声 (harmony) と呼ぶ。音楽理論には楽曲中の和音の機能が明確になるように、和音の接続に関する様々な指針が記述されている²。表 1 に典型的な古典西洋音楽における和音進行の例を示す。例えば V の和音から 4 度 (半音 5 つ) 上昇して I の和音に到達する和音進行はドミナント進行と呼ばれ、強い指向性を持つ。到達時に強い解決感を知覚させる性質があるため、調性音楽では定型的に多用される³。その一方で和音進行の指向性には非対称性があり、I の和

¹本稿では半音 5 つを 4 度、半音 7 つを 5 度と表記するが、音楽理論では歴史的な事情からそれぞれ完全 4 度、完全 5 度と表記されることも多い。

²古典的な音楽理論では、独立した複数の声部 (音) が逐次的に和音を構成するときの声部の動きについて規定されている。

³調性音楽ではより強い解決感を得るため、属和音に代わって 7 度音を付加してより音響的に

音を中心とする体系化に寄与している．なお，ナポリの六の和音⁴などの特殊な三和音を除いて，音楽理論では音階外のピッチクラスを根音として持つ和音に関する言及は少ない．

実際の楽曲の聴取では，楽曲の進行に伴って調性が内的に確立され，和音の機能が明確化していく．このような過程を調性的体制化と呼ぶ．音楽理論による解釈は分析上の解釈であるため，例えば曲頭の和音はそれ自体を聴取しただけでは楽曲の中の位置づけが分からず，機能や調性を同定することはできない．また，実際の音楽に含まれる和音の機能は，多義的に解釈可能な場合が多くある．例えば曲中で調性が変化する転調や，一時的に他の調からの和音を借用する一時的転調を含む箇所では，音階外の和音が用いられることになるため，多義性が生じやすい．

表 1 音楽理論上続きやすい和音進行

Piston (1948)[7] p.17 の表より改変．

現在の和音	よく続く	ときどき続く	ほとんど続かない
I	IV, V	VI	II, III
II	V	VI	I, III, IV
III	VI	IV	II, V
IV	V	I, II	III, VI
V	I	VI, IV	III, II
VI	II, V	III, IV	I
VII *	III		

(* VII を根音に持つことは滅多にない)

も緊張感を高めた属七和音を用いることが多い．また古典音楽理論では，和音を構成する声部の音高の選択についても，音響的側面及び前後の和音の持つ機能に基づいて決められるよう，厳格に規定されている．

⁴II の和音，ただし多くの場合最低音として IV の音を用いる．長音階の楽曲において音階外のピッチクラスを根音に持ちながら調性感を損なわない特殊な例として知られている．

2.2 和音進行における音楽的期待の心理学的研究

和音進行における音楽的期待は、特に 1980 年代以降、音楽理論から導かれる調や和音の機能との対比の元で数多くの実験心理学的研究が行われてきた。

先行する音楽的文脈と期待の関係を明らかにする手法として、Krumhansl らのグループによって提唱されたプローブ音法が挙げられる [8]。Krumhansl, Bharucha, & Kessler (1982)[9] は、先行して上昇または下降するスケール音列を呈示したのち、2つの和音が連鎖する刺激を聴いたときに生じる主観的な適切度のパターンについて調べた。多次元尺度構成法及び階層クラスタリングを用いて分析したところ、音楽理論上重要とされる音程の進行について高く評価される傾向がみられた。この研究では、先に呈示したスケール音列による影響については大きな差がなく、詳細な検討はなされていない。また3つの調と関係が深いとされる13種類の和音を組み合わせた刺激を用いており、狭い範囲の連鎖和音しか対象とされていない。しかし、先行して調を同定できる刺激を呈示した後、続いて単音や和音等のプローブ音について適切度や終始感等の主観量を問う方法論は、のちに様々な拡張がなされ調性知覚を論じる代表的な手法へと発展した [1, 9, 10]。プローブ音法から得られる、ある調を与えられたときに各ピッチクラスの単音や和音に対して生じるヒトの主観的な評価値の分布はキープロファイルと呼ばれ、後述する楽曲の統計的な研究にも影響を与えた。

一方で、判断課題に対する反応時間から期待の大きさを見積る実験も数多く行われてきた。Bharucha, & Stoeckig (1986) は2つの和音からなる連鎖和音を呈示し、2番目の和音について通常の音高 (in-tune 条件) か、構成音の音高がわずかにずれているか (out-of-tune 条件) を判断する課題を実施した [11]。その結果、音楽理論上重要な音程の和音進行のほうが、反応時間が有意に短くなることを示した。期待が大きい刺激ほどそれに対する反応時間が短くなると考えられるため、音楽理論上の重要度と知覚上の期待の大きさと的一致を示唆する結果であると言える。さらに、そのような期待されやすい和音進行は、前後の和音に共通する構成音が含まれる場合が多いが、共通する構成音が存在しない和音進行で比較しても同様の効果が見られることも報告されている [12]。

このような局所的な和音間の音程に応じて生じる期待に加え、より長い連鎖数

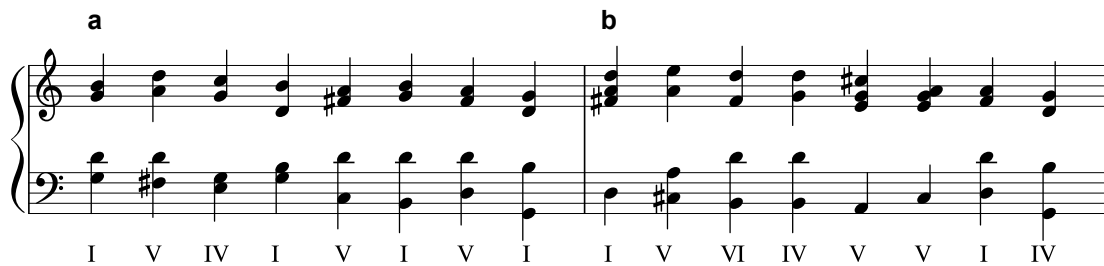


図 2 音楽理論を前提とした連鎖和音刺激の譜例

Bigand ら (1997)[13] Figure.1(p.1101) より抜粋・改変。譜面下段に音楽理論上想定される調性に基づく和音名を記す。(a), (b) 共に最後の 2 つの和音は物理的に等しいが、それ以前の文脈によって異なる調に導かれるため、和音の持つ機能が異なるとされる。

の連鎖和音を用いた場合には和音の文脈に応じた大局的な期待も生じる。Bigand, & Pineau (1997) は 8 つの和音からなる連鎖和音刺激を用いて、最後の和音について主観的に適切かどうかを回答する課題や in-tune または out-of-tune かの判断を行う課題等を実施した [13]。このとき、後ろの 2 つの和音は同一とし、その前の 6 つの和音を変更することで音楽理論上の調が異なる条件を用意した。図 2 に示す譜例では、(a) と (b) 共に最後の 2 つの和音が D メジャーコードから G メジャーコードへ接続されるが、それ以前の連鎖和音がそれぞれ G メジャーキー、D メジャーキーの調を有するため、和音のもつ調性上の機能が異なる。実験の結果、先行する和音の文脈によって導かれる調において、より頻出する和音進行のほうが稀な進行と比べ適切度が高く、判断の反応速度が速くなることが示された。

こうした和音進行における期待を説明するモデルとして、活性化拡散モデルが提唱されている [14]。活性化拡散モデルは意味記憶のモデルとしてよく知られている [15]。このモデルでは、音高、和音、調の 3 つの階層を想定しており、各階層を構成するノードは隣り合う階層の互いに関連のノードと結合されているものとする。呈示された音から生じた活性が、モデル内のネットワークの近い関係にあるノードを次々活性化させながら平衡状態に達するまで拡散する。さらなる観測を繰り返すことで、観測した和音系列に当てはまりのよい調がだんだんと選択されていく。ただしこのモデルが機能するためには、和音と調の階層を繋ぐ接続パターンが予め音楽理論上親和性の高いノード同志に限られている必要がある。

このような接続のパターンは、個人の音楽的な経験によって形成され则认为られている。Tillmann, Bharucha, & Bigand (2000)[16] は自己組織化マップを用いて活性化拡散モデルの獲得過程について検証し、調性音楽へのばく露を前提とすることで同様の接続パターンが得られることを示唆した。活性化拡散モデルは自動的で素早い親近調の内的な確立 [17] や和音の順番に対する頑健性 [18] をうまく説明できる。

近年では、和音の音符の時間的構造による影響 [19] や、連鎖和音中の位置による影響 [20] などについて検討されている。

また、非侵襲型脳活動計測装置を用いた研究報告も盛んである。Koelsch et al. (2000)[2] は調性から逸脱する和音に対して、左半球前頭において 200 ms 程度の潜時に生じる陰性の電位変化を報告した。ERAN (Early Right Anterior Negativity) と呼ばれるこの電位変化は、近い潜時に生じる予想外の刺激の検知によるミスマッチ陰性電位 [21] とは別種のものであるとされる [22, 23]。ERAN の潜時は一般に認知的な処理と呼ばれる電位変化と比べると短く、内的な調性への反応が無意識下で行われていることと整合性がある。実際、音刺激に対して注意を向けない条件下においても、活動が観測されることが報告されている [24]。脳磁計を用いた実験でも同様の活動が報告され、電流源は下前頭回のブローカ野及び対側の領域（ブロードマン 44 野）に由来するとされる [25, 26]。fMRI を用いた研究でも同領域の活動が確認されている [27, 28]。ブローカ野は感覚入力のもダリティ依存性が低い高次連合野のひとつであり、言語理解の文法処理に携わる [29, 30]。和音進行のパターンは調性音楽において文法的意味合いを持つことから、言語と音楽の共通性に関する議論も盛んに行われている [31, 32, 33]。

2.3 楽曲構造の情報学的研究

計算機環境の発達に伴い、楽曲の構造を情報論的な枠組みから分析する研究も並行して発展してきた [34, 35]。音楽情報科学と呼ばれるこの分野では、楽曲の構造の統計的性質をモデル化することで、楽曲の合成、ユーザーの趣向に応じた楽曲推薦など、工学的応用も視野に入れた様々な研究が行われている。

調の自動推定は最も古典的な研究トピックのひとつである。楽音の出現頻度が調によって異なることを利用することで、調を推定することができる。Krumhansl-Schmuckler モデル [1] は、キープロファイル [10] と楽曲を構成する音の各ピッチクラスごとの頻度分布の相関（または差の大きさ）を元に調を推定するモデルである。Temperley(2007) はさらに調の構造を考慮して、ベイズ理論に基づく調推定を提唱した [6]。予め音楽理論の専門家が機能和声のラベルを付与した楽曲データベースを用いて調ごとの構成音の出現頻度分布を求め、推定対象の楽曲の頻度分布に対する当てはまりを評価することで、高い精度で調を推定することができる。一方で、ヒトによる推定の結果とはかい離が見られることも報告されている [36]。

和音進行のパターンに着目した研究では、マルコフ過程によるモデル化が行われてきた [37, 38]。マルコフモデルは音楽理論における和音進行のルールを数学的に記述する基本的な方法であり、楽曲データベース内のパターンを分析する方法としても用いられる [39]。

近年は音楽の文法的な構造を考慮して、階層的な木構造を考える生成音楽理論 (GTTM; Generative Theory of Tonal Music, [40]) による楽曲分析が報告されている [41, 42]。さらに生成音楽理論をヒトの知覚モデルとして捉えなおした IDyOM (Information Dynamics Of Music, [43]) を用いた心理実験や脳活動計測結果の検証など、楽曲構造とヒトの知覚を結び付ける議論が盛んである。しかし、モデルの内部パラメータをアドホックに設定する必要があるなど、複雑すぎる故に知覚実験との比較の評価は難しい。

2.4 従来研究で残された課題

和音進行における音楽的期待に関連する様々な側面について研究が進められてきた。これらの研究で得られてきた知見はヒトが楽曲の調性的な構造を、ヒトが心理的な表象として同様に捉えていることを前提とした上で得られた結果である⁵。この前提を保証するために、調による影響を検討した先行研究では、音楽理論上和音の機能について多義性を生まない音楽刺激を用いてきた。例えば図2に示した Bigand, & Pineau (1997)[13] の2つの実験刺激は、音楽理論上調が明確に定義できるように音階外のピッチクラスを含まない8連鎖和音を採用している。もし図2(a)の譜例における末尾の3つの和音のみを取り出した3連鎖和音を用いた場合、Gメジャーキーと類推されるが音階上IVの音に対応するピッチクラスであるCの音を含まないため断定することはできない。そのため、V-Iの標識に基づく和音進行の分析は音楽理論上も保証が難しい。図2(b)の譜例における末尾の3連鎖和音を用いた場合、構成音はDメジャーの音階を満たすが、Vの和音で始まるため調性音楽として特殊な例と見なされる。そのため、やはりI-IVの標識に対して議論の余地が残る。一般に3連鎖和音など短い連鎖和音は、調性的体制化が進む段階にあるとされ[10]、調を一意に標識できないことが多い。このような条件下では和音の機能について明に条件分けを行うことは難しいため、知覚パターンについて詳細に研究されてこなかった。同様の理由で、調性的な標識自体が存在しない非調性的な連鎖和音の組み合わせについても、調性からの逸脱刺激としての議論[24, 44]以外では避けられてきた。

ただし音楽理論自体、理論的基盤を持たない経験的な知見の集積に過ぎないことに注意する必要がある。これまでの多くの研究結果は、調性音楽を特徴付ける和音進行に対して音楽的期待が高くなる、という解釈で一貫して説明できる。しかし、期待の知覚は音楽理論で定義される調と依存関係にあるとされるため、音楽理論に依存したパラダイムの中で期待の知覚メカニズムの議論をすること自体が危うい[45]。近年の音楽情報科学における楽曲の構造分析を利用した知覚研究

⁵調の成立を前提とする必要のない2連鎖和音の研究においても、暗に調性面で偏った刺激を用いている場合がある。例えば、Krumhanslら(1982)[9]による研究では、一般的な長音階内で構成される和音の組み合わせについて検討を行っている。この場合音階の制約自体が調に関する情報を含んでいるとされるため、実験結果の一般性は保証されない。

も、実験刺激の持つ統計的性質を実験条件の標識として用いることになるため、同様の議論を抜け出せない。

一方でこれまでの和音進行における期待に関する研究結果は、局所的な音程変化の効果と大局的な文脈の効果の2種類の影響を受ける [46] ことを示唆していることから、多変数による複雑な情報処理を背景としていると考えられる。そのため、音楽理論に代わる新たな手法の導入が必要である。

2.5 計算論的アプローチ

ここでは、知覚における情報の流れを計算モデルとして記述する、計算論的アプローチと呼ばれる方法論に注目した。計算論的アプローチは、ヒトを入力（刺激）に対して出力（知覚的応答や行動選択）を返すシステムと捉え、同等の性能を持つ計算モデルの設計を試みることで、システムの中身を明らかにする。この方法論のコンセプトは、Marr(1982)の提唱した知覚システムを理解するために計算理論 (computational model)、表現とアルゴリズム (representation and algorithm)、神経細胞上における実装 (hardware implementation) の3つの階層 [47] を縦断的に扱うことにある。

複雑な認知現象に対する計算論的アプローチは、Montague, Dayan, & Sejnowski (1996)[48] や Schultz, Dayan, & Montague (1997)[49] によるサルの神経生理学的研究が重要な起点となった。彼らはドーパミンニューロンの発火パターンが、TD(Temporal Difference) 学習と呼ばれる強化学習の報酬予測誤差に対応して説明できることを示した。以降ヒトの研究においても、核磁気共鳴画像法 (fMRI; functional magnetic resonance imaging) で得られた脳賦活データに対して回帰分析を適用し、モデルに対応する計算処理を担う脳部位を同定するという方法論が確立されてきた [50, 51]。計算論的アプローチは時間的に変化する直接観測できない内的な変数（隠れ変数）を扱う、複雑な意思決定 [52] や推論 [53, 54, 55]、行動学習 [56] などの高次脳機能の研究分野において大きな成果を上げてきた [57, 58]。

本研究で扱う音楽的期待も、同様に複雑な情報処理に基づくことが示唆されている [14]。楽曲構造分析において情報論的なモデリングは一定の成功を収めており [6]、音楽理論とも親和性が高いと言える。このため、音楽的期待に関連する行動実験からそのデータの持つ性質を近似できるモデルを組み立てるという方法論は、背景メカニズムに迫る上で有用なツールになりうると考えられる。

2.6 本研究の指針

本研究では行動実験に基づく計算論的アプローチにより，和音進行に対する音楽的期待のメカニズムの解明を目指す。

期待のモデル化に当たって，本研究ではいくつかの点で従来の計算論的アプローチの研究と異なる工夫を行っている．一つ目は行動データに基づくモデルの設計方法である。

先述のように計算論的アプローチは，動物の神経生理計測で得られた知見と学習モデル [59] に含まれる変数の関係を巡る議論 [48, 49] から発展してきた．ヒトの高次脳機能を扱った研究では，工学の分野で提唱された理論を下地としてトップダウン的にモデルを組み立てている [51]．例えば Badre ら (2012) [54] の研究では，探索的な意思決定に関する情報論的に妥当ないくつかのモデルを実験参加者の行動にフィッティングする形で推定し，実験参加者の行動をどの程度再現できているかを評価している．さらに彼らは推定したモデルを用いて fMRI 信号を回帰分析することで，相関する活動を示す脳部位を特定し，先行研究で得られてきた脳部位の機能に関する知見との関係からモデルの妥当性を論じている．しかし，音楽的期待は理論的基盤自体が乏しい．また，これまで神経生理学的な知見も得られていないため，同様のトップダウンな形でのモデルの構築は困難である．そこで本研究では，まず心理実験から得られた行動データについて明確なモデルを仮定しない形で統計的分析を行い，行動に含まれる特徴の抽出を行った．このようなボトムアップな分析を先行することで，モデルに含まれるべき変数の関係についての情報を得た．その結果に基づいてモデルの候補を組み上げる，セミ・トップダウンな戦略を採用し，モデルの理論的な基盤が乏しい条件において計算論的アプローチによる分析を可能とした⁶．

二点目はモデリングに用いるデータの問題である．通常確率モデルを用いたモデリングでは，多数の標本データが存在する条件下で行われる．一方，行動実験

⁶実験心理学において，モデリングは実験結果の最終的な解釈として用いられる（例えば，[14]の研究等）．そのため，神経科学的な知見に基づかないトップダウンなモデリングに基づく研究は受け入れられにくい．本研究におけるセミ・トップダウンな戦略は，従来のボトムアップな方法論によって導かれた知見に基づきつつも，それだけでは説明不可能な部分について最小限のモデリングを行うことで，方法論的な批判に応える側面もある．

から得られる標本データ数は、時間的な制約を受けて少数に限られる。また、データの取得方法についても、特定のモデルに対して有利にならないように注意する必要がある。そのため、本研究では二値強制選択ではなく評価尺度を採用し、少ないデータ数で実験参加者の回答から期待の傾向を取得できるようにした。さらに、可能な範囲で全探索的に条件を設定した。

三点目はモデルのパラメータ推定と性能比較の方法である。従来の神経科学においては、モデルのパラメータを実験データをよりよく説明できるようにハンドチューニングで設定することがしばしば行われてきた。本研究ではモデル選択を含めて議論するため、実験参加者の行動データから直接パラメータ推定を行った。また複雑なモデルほど実験データをよく説明できてしまう現象を避けるため、本研究では標本数が限定される個人ごとの分析ではモデルの学習に用いていない未知データに対する汎化性能評価を、全参加者のデータを用いたグループ分析では交差検定及びパラメータ数の効果を補正した情報量規準による比較を採用した。

心理実験の条件設定は、上述の形でモデルの設計と評価が可能となるように心掛けた。従来の多くの心理学研究で用いられる対照実験のデザインは、実験条件とコントロール条件を比較する統計的仮説検定に特化しており、本研究のようなセミ・トップダウンなモデル構築には不相応である。また、モデル内の変数の数が多いほど候補となるモデルも増加してしまうため、理論的基盤のない現段階では長い連鎖和音を用いることは現実的ではない。その一方で従来の音楽理論に基づく研究では取り扱いが困難であった調性が曖昧な条件において、計算論的アプローチは拘束を受けないメリットがある。以上のことから、本研究では最長4連鎖までの長3和音で構成された短い連鎖和音における期待に着目した。その上で、可能な限り全探索的な形で実験条件を設定した。

この実験の長3和音の組み合わせによる条件設定では、2連鎖条件に含まれる12種類の音程のうち、音楽理論上の音階内に収まるものは長2度、4度、5度の3種類となる。また3連鎖条件では、同一和音の連続を除く132種類の音程の組み合わせのうち、音階内に収まるものは6種類である⁷。和音の根音にのみ着目し

⁷音楽理論上、長音階におけるI、IV、Vの和音の組み合わせから自己遷移を除けばよいため、 ${}_3P_2 = 6$ 通りである。ただし、特殊な事例として許容されるナポリの3和音を含む進行（♯II-V）や、ドミナントの機能を持つ和音（V）に対してドミナント進行するダブル・ドミナントと呼ば

て、その他の構成音が音階外になることを無視した場合でも、 ${}_7P_2 = 42$ 種類である。すなわち、本研究で用いられる実験刺激の多くは、一般的な調性音楽で用いられる和音進行ではない。

れる和音進行（II-V）など、実際の調性音楽では様々な拡張的な解釈を行うことがある。

第3章 短い連鎖和音における和音進行への音楽的期待

3.1 目的

短い連鎖和音は調性的体制化が進む段階にあるとされ [10], 音楽的期待と調性の関係を明らかにする上で重要な条件であると考えられる. しかし, 音楽理論上調性が一意に同定できない組み合わせが多いため, 最も単純な2連鎖和音を除けばあまり研究されてこなかった.

本章では続く第4章での計算論的アプローチによる分析を念頭に, 短い連鎖和音の呈示において生じる和音進行の期待が連鎖和音のどの変数にどのような影響を受けているかを明らかにすることを目的とする.

実験課題は, Krumhansl, Bharucha, & Kessler (1982)[9] や Bigand, & Pineau (1997)[13] による先行研究を元に発展させた, 評価尺度による主観的適切度の判断課題を採用した. Bigand, & Pineau (1997) の研究によれば, 強く期待される和音進行に対して主観的適切度の評価が高くなる傾向がある.

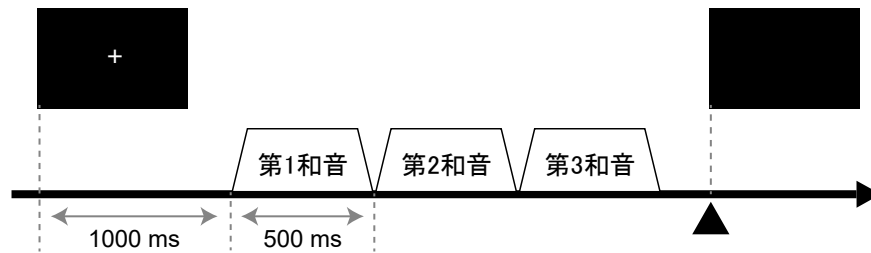


図 3 実験のフロー

例として 3 連鎖条件における刺激呈示と回答のフローを模式的に表す．黒三角形は実験参加者のボタン押し回答のタイミングを表す．

3.2 手法

3.2.1 心理実験

実験参加者は正常な聴覚を有する成人 15 名とし（男性 9 名，女性 6 名，年齢 24.9 ± 4.0 歳，日本の義務教育以外の音楽経験 10.1 ± 5.7 年），その全員について九州大学大学院芸術工学研究院の承認を受けた倫理基準を満たすインフォームドコンセントを得た．事前に実験内容を説明し，実験手続きについて十分理解していることを確認した．

実験参加者に 2 つから 4 つの和音が連続して再生される連鎖和音刺激を呈示し，その最後の和音に対する文脈上の適切度の大きさについて回答を求めた．回答は一行に並んだ 9 段階のボタン押しとし，「どちらでもない」を指す「5」を境に左端の「1」に迫るほど「より適切でない」，右端の「9」に迫るほど「より適切である」，とした．このとき，最後の和音の呈示からなるだけ早く直観的に回答するように指示した．参加者の前にはボタンの他にディスプレイを設置し，最初の和音が鳴りはじめる 1 秒前からボタン押しを検知するまでの間白色の十字の固視点を呈示した（図 3）．

呈示する和音は，純音によって構成された 12 のピッチクラス全てを根音とする長三和音から選択した．和音の構成音は平均律 $A4=440$ Hz でチューニングした．和音の重ね方による回答への影響を軽減するため，各和音は G3-F \sharp 6 の範囲にお

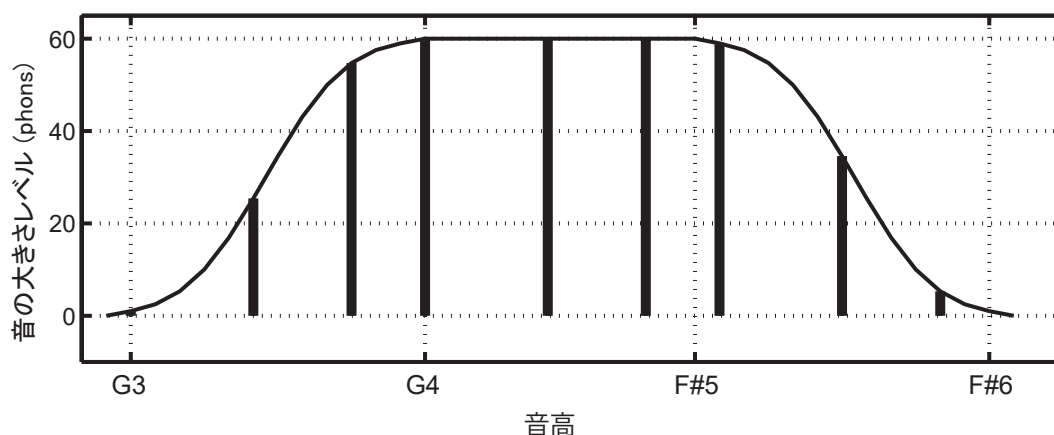


図 4 実験に用いた和音刺激の例

G メジャーコードの構成音の例を模式的に示す。本実験では Krumhansl, Bharucha, & Kessler (1982)[9] で用いられたシェパードトーン [60] の和音を参考に MATLAB(MathWorks 社) を用いて刺激を合成した。この例では G3 から D6 までの計 9 つの純音にエンベロープを掛け合わせた形になっている。ただし、音の大きさレベルを算出するための等ラウドネスレベル曲線は、2003 年に改訂された国際規格 (ISO 226:2003) に従った。

いて和音を構成するピッチクラス全てを鳴らし、さらに G4 から F#5 の範囲の音量は等しく、その範囲から離れるにつれ減衰するように合成した [60] (図 4 参照)。合成された刺激音はアンプ (STAX SRM-323S) に接続したヘッドフォン (STAX SR-307) を用いて両耳へおよそ 60 dB (A 特性) で呈示した。刺激呈示及びボタン押しデータの収集は MATLAB(MathWorks 社) の Cogent toolbox (Wellcome Department of Imaging Neuroscience: <http://www.vislab.ucl.ac.uk/Cogent>) を用いた。

実験は 3 種類の連鎖数ごとにセッションを分け、セッションの順番は連鎖数が 2 つ、3 つ、4 つの順とした。それぞれのセッションの前に、ランダムに選んだ 20 試行からなる練習セッションを実施して参加者の課題内容に対する理解を確認した。

2 連鎖条件で用いた連鎖和音は、12 種類の長和音の取りうる全ての組み合わせを含めた (144 試行)。その他の条件では、全ての和音の組み合わせについて実

験することが実験時間の制約上現実的ではないため、和音同士の音程関係に着目して試行数を減らした⁸。3連鎖条件では和音間の音程の取りうる全ての組み合わせ（144 試行）とした。4連鎖条件では、直前の和音との音程が特定の音階を含まないように、オクターブ内で対称性を持つ音程の組み合わせの中から長2度、完全4度、完全5度、短7度のいずれかとなるパターン（64 試行）に限定した。第一和音はランダムに選択した。いずれの条件も、連鎖和音の呈示順序はランダムとした。

和音の長さは一つの和音あたり 0.5 秒とした。ボタン押しから次の試行までの時間間隔は 1.5 秒から 3 秒の間とし、試行ごとに 0.3 秒刻みでランダムに選択した。セッション間の休憩時間は実験参加者の裁量で決められるようにした。実験時間は休憩を含めておよそ 1 時間であった。

3.2.2 分析

実験で得られた主観的な適切度を degree of relatedness (DOR) と呼ぶ。DOR がどのような情報に影響を受けているのかを調べるため、統計的検定及び階層クラスタリングを用いて分析した。分析には全参加者間で平均した DOR を用いた。

⁸従来研究では、和音進行の期待に対し音高が寄与しないことは当然の前提として扱われている。本研究でも事前の予備的な計測でこのことを確認しており、2連鎖条件以外では音程による影響についてのみ検討することとした。

3.3 結果

各連鎖和音に対する DOR の実験参加者間平均値を図 5 に，標準偏差を図 6 に示す．

3.3.1 統計的分析

2 連鎖条件の DOR について，2 つの和音の音高と両者間の音程の 3 つの因子が与える効果を調べた (図 7)．その結果，音程の効果が有意であった一方 (Kruskal-Wallis 検定 $\chi^2(11) = 100.8; p < 0.001$)，音高の効果は有意でなかった (第 1 和音： $\chi^2(11) = 8.0; p = 0.71$ ，第 2 和音： $\chi^2(11) = 4.7; p = 0.95$)．この結果は，DOR が和音の音高ではなく音程に依存して決まっていることを示している．

3 連鎖条件の DOR には，標的和音である第 3 和音に対して第 1 和音からの音程と第 2 和音からの音程の 2 つの因子が存在する (図 8)．Friedman 検定を用いて各因子の主効果を調べたところ，第 1 和音，第 2 和音共に主効果が有意であった (第 1 和音： $\chi^2(11) = 417; p < 0.001$ ，第 2 和音： $\chi^2(11) = 85.1; p < 0.001$)．この結果は DOR が隣り合う第 2 和音だけでなく，さらに一時刻前の第 1 和音からの音程にも影響を受けていることを示唆している．

3 連鎖条件における第 1 和音からの音程による影響のパターンについて調べるため，3 連鎖条件の DOR から対応する一時刻前の音程の 2 連鎖条件の DOR を差し引いた．その結果を図 9 に示す．第 1 和音と第 3 和音が一致もしくは音程が 5 度 (半音 7 つ) の関係にある刺激に対する差分 DOR と，短 2 度 (半音 1 つ) もしくは短 6 度 (半音 8 つ) の関係にあるときの差分 DOR との間に有意差が認められた (Friedman's test, $\chi^2(11) = 85.1; p < 0.001$; mean-rank post-hoc test with Bonferroni correction, $N = 12, p < 0.00076$)．

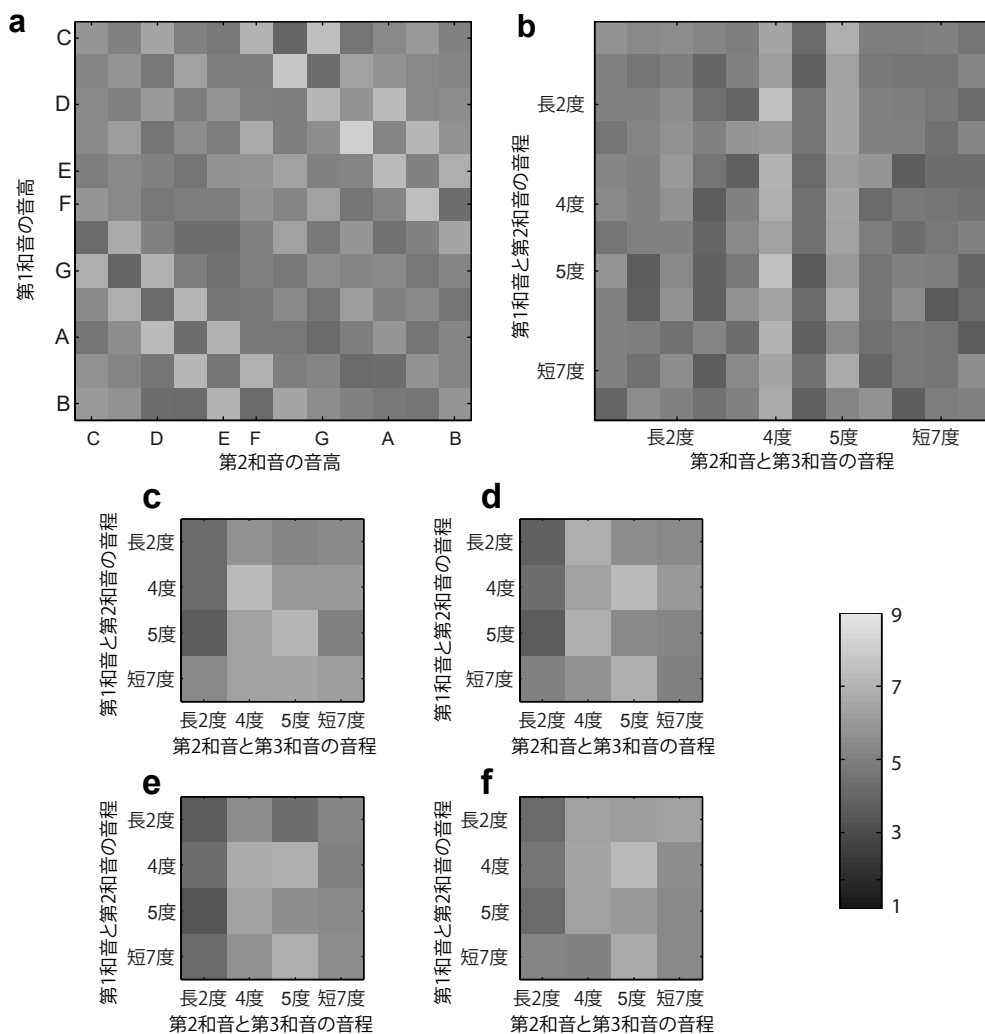


図 5 各連鎖和音に対する全参加者の平均 DOR

(a) 2連鎖条件における DOR. 縦軸方向に第 1 和音, 横軸方向に第 2 和音の根音の高さを取り, グレースケールの濃淡で値を示した. (b) 3連鎖条件の DOR. 縦軸方向に第 1 和音と第 2 和音間の音程, 横軸方向に第 2 和音と第 3 和音間の音程を取った. (c-f) 4連鎖条件の DOR. 第 3 和音と第 4 和音間の音程の種類ごとに 4 つのグラフに分けて示す (c: 長 2 度, d: 4 度, e: 5 度, f: 短 7 度). それぞれ縦軸方向に第 1 和音と第 2 和音間の音程, 横軸方向に第 2 和音と第 3 和音間の音程を取った.

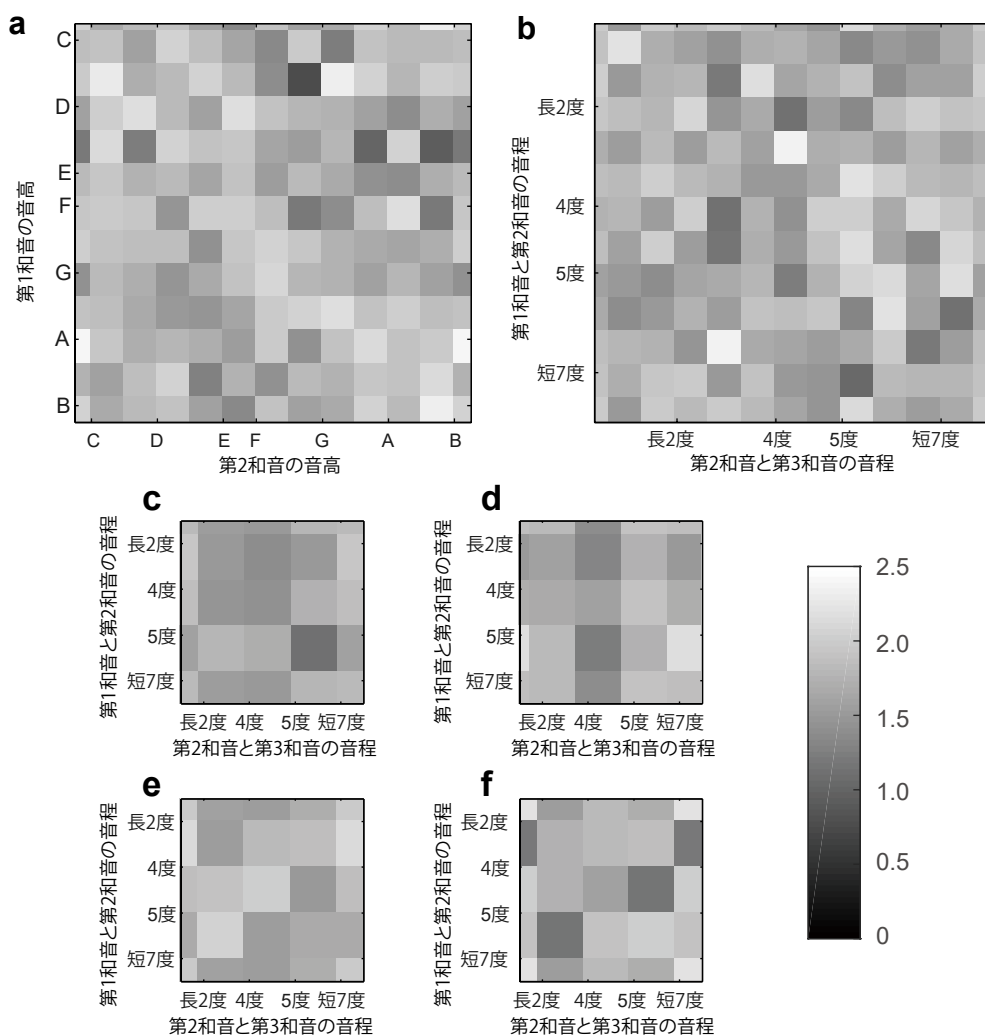


図 6 各連鎖和音に対する DOR の参加者間の標準偏差

(a) 2 連鎖条件における DOR の標準偏差, (b) 3 連鎖条件の DOR の標準偏差, (c-f) 4 連鎖条件の DOR の標準偏差をそれぞれ示す. 図の軸は図 5 と同様とし, グレースケールの濃淡で値を表す.

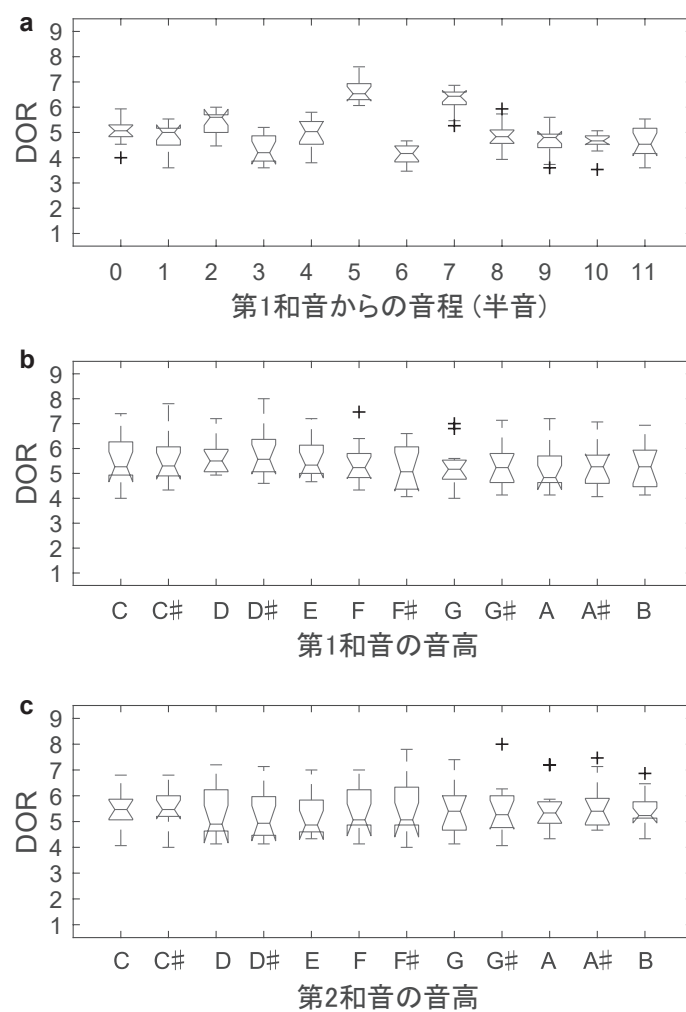


図 7 2連鎖条件における DOR の分布

(a) 音程, (b) 第 1 和音の音高, (c) 第 2 和音の音高にそれぞれ着目したときの DOR の分布を箱ひげ図で表す.

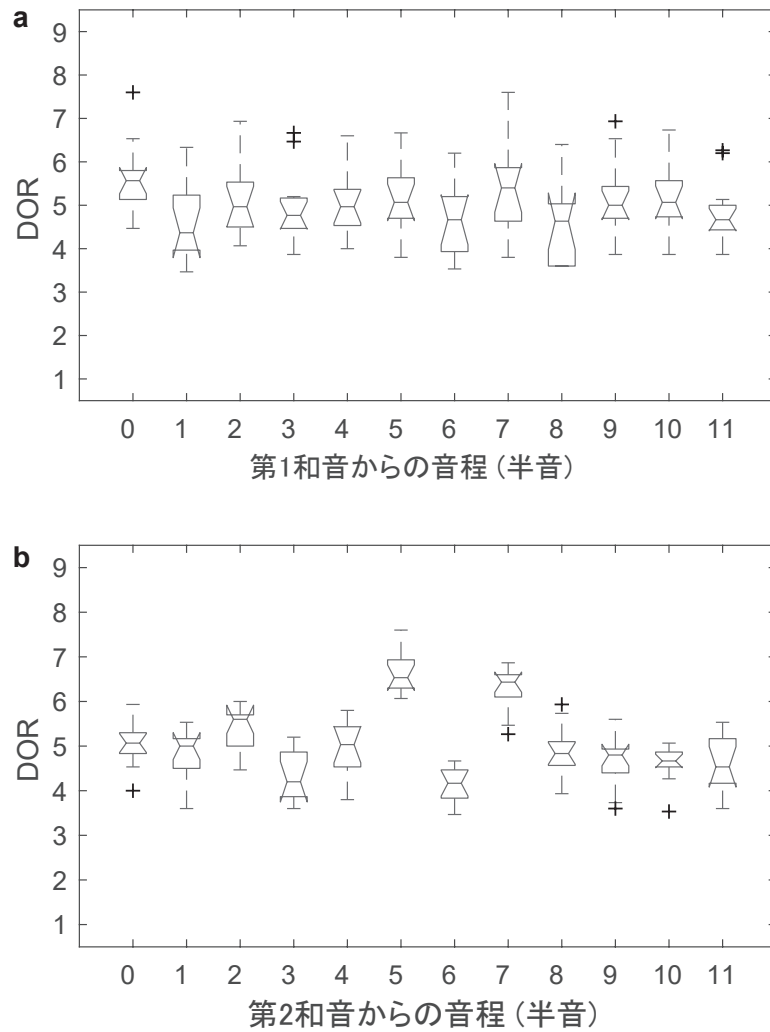


図 8 3連鎖条件における DOR の分布
 (a) 第 1 和音と第 3 和音の音程, (b) 第 2 和音と第 3 和音の音程それぞれに着目したときの DOR の分布を箱ひげ図で表す.

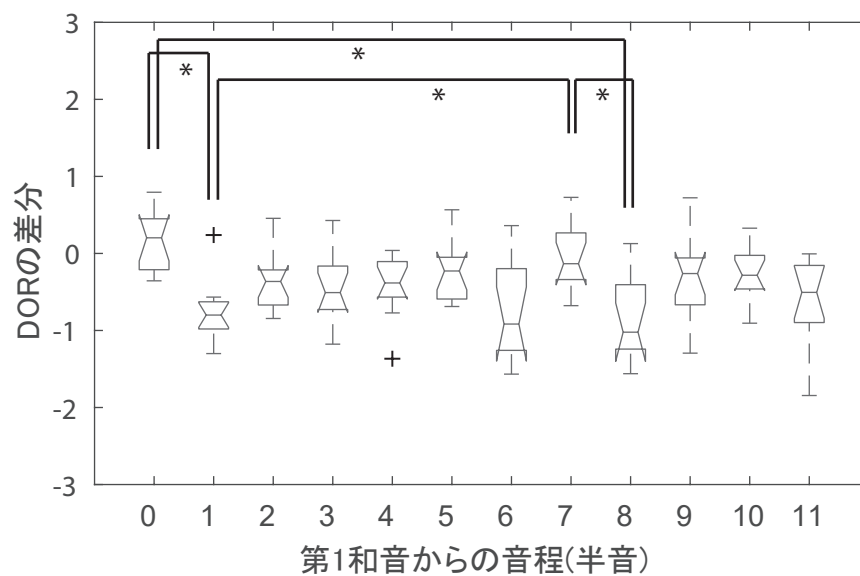


図 9 3連鎖条件と2連鎖条件のDORの差分

正の値ほど3連鎖条件のほうが2連鎖条件よりも大きい評価値であったことを示す．下位検定で有意差の認められた組み合わせをアスタリスクで表す (mean-rank post-hoc test with Bonferroni correction, $N = 12, p < 0.00076$).

3.3.2 多変量解析

和音間の音程による影響のパターンについて調べるため、4連鎖条件を含めた DOR の類似傾向について分析した⁹。本実験で用いた連鎖和音には、DOR を回答した最後の和音に対して6通りの和音位置の組み合わせ（2連鎖条件は1通り、3連鎖条件は2通り、4連鎖条件は3通り）が含まれる。それぞれの組み合わせが DOR に与える効果の類似度を、多変量解析により確かめた。まず全ての和音の組み合わせに共通して含まれる4種類の音程に着目し、それぞれが呈示されたあとの DOR の平均値を求めた（図 10(a)）。次に4種類の音程に対する平均 DOR の大きさを軸にとった4次元空間におけるユークリッド距離行列を算出し、非計量多次元尺度構成法 [61] を適用した（図 10(b)）。その結果、第2和音と第3和音、第3和音と第4和音の DOR の傾向が近いことが分かった。

また4種類の音程に対する平均 DOR（図 10(a)）に対し、最長距離法に基づく階層クラスタリング [62] を適用した。その結果、時間的に連続する音程による効果と、それ以外の大きく2種類のクラスタに分類できた（図 11）。この結果は、一時刻前の和音から受ける影響が、和音の位置に依らず似ていることを意味している。その中でも特に、第2和音と第3和音、第3和音と第4和音の組み合わせにおける DOR の傾向が近いことが示された。

⁹4連鎖条件は4種類の音程（長2度、4度、5度、短7度）の組み合わせで構成された刺激のみを用いており、音程の組み合わせが限定されるため、他の条件と同様な統計的検定を適用することはできない。

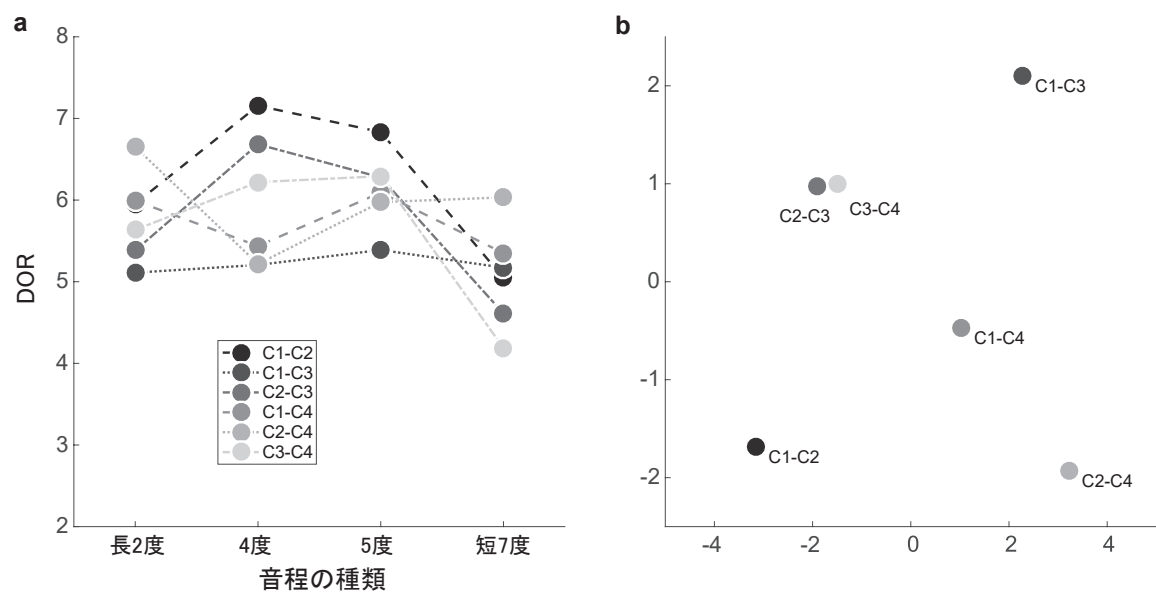


図 10 多次元尺度構成法による，DOR のパターン分析

a) 6通りの和音位置の組み合わせにおける平均DORのプロット．共通して含まれる4種類の音程を横軸とする．図中のレジェンドにおける C_m-C_n は m 番目の和音と n 番目の和音の組み合わせを指す．類似度を直感的に理解しやすくするため，同じ和音位置の組み合わせを点線で結んでいる．b) 多次元尺度構成法によって得られたマップ (stress < 0.001)．円はそれぞれ和音位置の組み合わせに対応し，円同士の距離によってDORの類似度を表す．

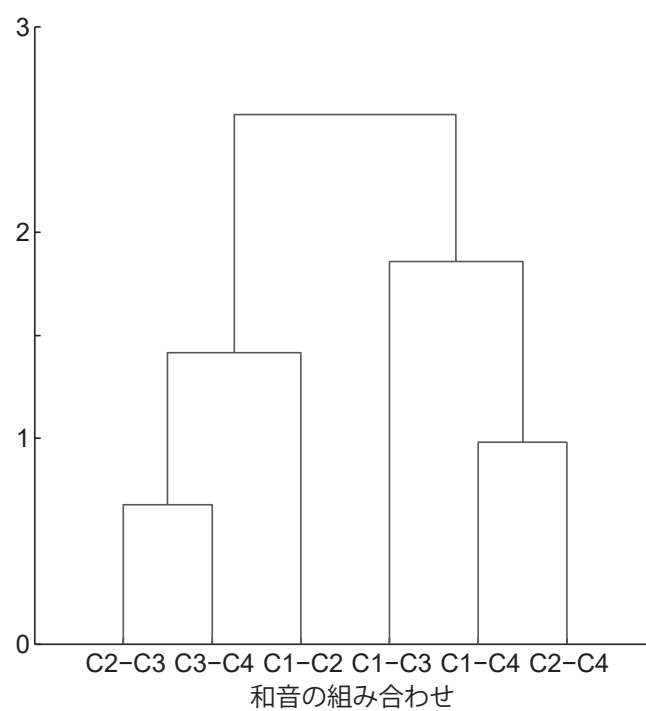


図 11 階層クラスタリングによる DOR のパターン分析
 パスの長さによって各音程が DOR に与える影響の類似度を表す。横軸の和音の組み合わせの表記は図 10 に従う。

3.4 考察

本章では和音間の音程関係に着目し、連鎖和音における主観的適切度 (DOR) について分析を行った。2 連鎖和音の条件では音高ではなく音程が DOR に寄与している、という従来からの知見を支持する結果を得た。

3 連鎖条件では、DOR に第 3 和音の直前にある第 2 和音だけでなく、隣接していない第 1 和音も影響を与えていることが示唆された。3 連鎖条件における第 1 和音の影響を、2 連鎖条件との差分によって見積った結果、第 1 和音と第 3 和音が一致、もしくは 5 度の関係にあるときに差が相対的に大きい値になる傾向があった。特に前者は物理的に等しい和音であり、実験参加者が 3 連鎖条件において内的に第 1 和音を保持していた可能性を強く示唆する。

一方で差分分析で有意差の認められた組み合わせに注目すると、差がやや正の傾向にある音程 (1 度と 5 度) と強く負に傾向にある音程 (短 2 度と短 6 度) が半音ずれるような規則性が見受けられる (図 9)。これは、音楽理論に基づく以下のように説明できる。第 1 和音を主和音とした場合、5 度の和音は属和音に相当する。もし第 1 和音を主和音と仮定すれば、調性を特徴付ける属和音に対しても期待が高くなるものと考えられる。また、第 1 和音を下屬和音 (IV の和音) と捉えたという解釈でも、期待が高くなることを説明できる。下屬和音を冒頭に置くことは、特に現在のポピュラー音楽では頻出する技法である。

2 つの多変量解析からは、隣接する和音からの音程が等しいときに DOR が類似する傾向が示された。この結果は、隣接する和音との音程が、短い連鎖和音の DOR において共通する形で寄与していることを示唆する。2 連鎖和音に先行してスケール音を呈示した先行研究 [9] でも、先行音の違いによる主観的適切度のパターンの差は見られておらず、音程に由来する局所的な期待 [63, 46] の存在を支持すると言える。また一方で、2 連鎖条件における DOR (図 10 および図 11 中の C1-C2) は、他と比べてやや異なる傾向が見られた。このことは第 1 和音が DOR に対し何らかの特別な寄与をしている可能性を支持する。

音楽理論上、短い連鎖和音の聴取は調性的体制化が進む途中段階にあるとされる。DOR の傾向は、第 1 和音が体制化の初期段階において調性的な中心、すなわち主和音としての役割を担っていることに起因すると解釈できる。しかし、和

音の連鎖に伴って体制化が進んでいく過程は，本分析では示すことができていない．調性的体制化を含む音楽的期待のメカニズムを理解するためには，期待に關与する音程等の変数間の関係を明らかにする必要がある．

なお，DOR の参加者間標準偏差は，連鎖和音によって違いが見られた（図 6）．このようなパターンは参加者間の応答に個人差がある可能性を示唆している（付録 B.1 における議論参照）．

第4章 音楽的期待の計算論的分析

4.1 目的

第3章で扱った古典的な分析手法から，和音進行に対する主観的適切度 (DOR) についていくつかの性質が明らかになった．1点目はDORが音高ではなく音程に依存して決まることである．2点目は連続する和音間の音程が同じとき，DORも類似する傾向があることである．3点目は，DORが評価対象である標的和音とは直接隣接していない和音からも影響を受けることである．しかし，影響する変数間の関係は依然明らかにできていない．

本章では，和音進行に対する音楽的期待が生じる内的な計算過程を明らかにすることを目的とする．上述のDORの3つの性質を満たすように，計算過程のモデル化を試みる．モデル化の指針は2.6節に従う．候補モデルを第3章で得られた行動データを用いて評価し，モデル選択を行う．モデルの評価は個人ごとと全参加者混合の2パターンで実施した．その上で得られたモデルと音楽理論を含む先行研究の知見と比較する．

従来研究では音楽的期待に調性の関与が指摘されており [13, 46]，計算モデルには調に関わる変数が含まれると考えられる．また第3章の分析において，第1和音と第3和音が同一のときにDORが大きい値を示していたことから，その調の形成に第1和音が強く寄与していると予想される．

4.2 方法

4.2.1 音楽的期待のモデリング

今、時刻 t ごとに 1 つの和音を逐次観測するものとする。このとき時刻 $t \geq 2$ において観測された和音 \mathbf{c} へ向かう進行に対する内的な適切度を、時刻 $t - 1$ までに得られた観測履歴の条件付確率を用いて以下のように表現する。

$$E(c_t = \mathbf{c}) = P(c_t = \mathbf{c} | c_{1:t-1}) \quad (1)$$

ここで c_t は時刻 t で観測された和音を表し、 $c_{1:t-1}$ は直前までに観測した連鎖和音を表す。連鎖和音は、平均律の 12 のピッチクラスを根音とする 12 種類の長和音の候補で構成される。

$$c_t \in \{\text{A major}, \text{A}\sharp \text{ major}, \text{B major}, \text{C major}, \dots, \text{G}\sharp \text{ major}\}. \quad (2)$$

第 3 章の行動分析の結果から、DOR は和音の音高ではなく和音間の音程によって決まることが示唆された。DOR の音程への依存は、音程の内的な基準を表す確率変数 X を導入することで表現できる。 X は全音高に対応する 12 個の状態を確率的に取り得る。ここでは便宜上 X を c_t と同じく和音として扱う。

$$X \in \{\text{A major}, \text{A}\sharp \text{ major}, \text{B major}, \text{C major}, \dots, \text{G}\sharp \text{ major}\}$$

式 (1) は X を導入することで次のようになる。

$$E(c_t = \mathbf{c}) = \sum_X P(c_t = \mathbf{c} | X, c_{1:t-1}) P(X) \quad (3)$$

また、一時刻前の和音から音程が同じとき DOR に与える効果が類似していることが示唆された。この効果は、尤度にマルコフ連鎖を導入することで表現できる。

$$E(c_t = \mathbf{c}) = \sum_X P(c_t = \mathbf{c} | X, c_{t-1}) P(X) \quad (4)$$

さらに、直接隣接していない和音からの影響もモデル化する必要がある。単純なマルコフモデルではこの効果を表現することができないため、更なる拡張を行う。

ベイズ更新モデル: Bayesian updating model マルコフモデルの拡張方策のひとつが、状態遷移確率の階層的な拡張である。即ち確率変数 X の分布を観測に応じて動的に更新するモデルが考えられる。時刻 $t \geq 2$ において X の各状態を取りうる最適な確率は、ベイズの定理とマルコフ性の仮定に基づいて逐次更新式の形で記述できる。

$$P(X|c_{1:t}) = P(X|c_t, c_{1:t-1}) \quad (5)$$

$$\propto P(c_t|X, c_{1:t-1})P(X|c_{1:t-1}) \quad (6)$$

$$= P(c_t|X, c_{t-1})P(X|c_{1:t-1}) \quad (7)$$

ここで、マルコフモデルの遷移行列に対応する尤度項 $P(c_t|X, c_{t-1})$ は、行列 \mathbf{T} で表現する。以後、本行列を期待行列と呼ぶ。期待行列の要素 $\mathbf{T}(i, j)$ は、 X が与えられた元で和音 i を観測したあと、和音 j を観測したときの期待の大きさを意味する。モデルの評価を行う第3章の実験では12種類の和音が用いられるため、 \mathbf{T} は144種類の組み合わせが含まれる。ここで各要素の和音は音高ではなく、確率変数 X からの音程(半音単位)でインデックスを付与する。 \tilde{c} を X からの音程に基づいて標識し直した和音とすると、各要素は以下のように表される。

$$\mathbf{T} = \begin{bmatrix} P(\tilde{c}_t = 0|\tilde{c}_{t-1} = 0) & P(\tilde{c}_t = 0|\tilde{c}_{t-1} = 1) & \dots & P(\tilde{c}_t = 0|\tilde{c}_{t-1} = 11) \\ P(\tilde{c}_t = 1|\tilde{c}_{t-1} = 0) & P(\tilde{c}_t = 1|\tilde{c}_{t-1} = 1) & \dots & P(\tilde{c}_t = 1|\tilde{c}_{t-1} = 11) \\ \vdots & \vdots & \ddots & \vdots \\ P(\tilde{c}_t = 11|\tilde{c}_{t-1} = 0) & P(\tilde{c}_t = 11|\tilde{c}_{t-1} = 1) & \dots & P(\tilde{c}_t = 11|\tilde{c}_{t-1} = 11) \end{bmatrix} \quad (8)$$

例えば、もし X が $A\sharp$ メジャーであった場合、 D メジャーコードは X の半音4つ上の和音となり、 $\tilde{c}_t = 4$ となる。 D メジャーコードを聴取したあとに $A\sharp$ メジャーコードを聴いたときの期待の大きさは $\mathbf{T}(1, 5)$ の要素として表現される。

参照和音 X の初期分布は、観測への重み付けの大きさを表すパラメータ κ を用いて以下の形で与えた。

$$P(X|c_1) = \begin{cases} \kappa & (X = c_1), \\ \frac{1-\kappa}{11} & (\text{otherwise}) \end{cases} \quad (9)$$

ここで $0 < \kappa \leq 1$ を満たす．このモデル化により，第 1 和音への重み付けの具合を評価することができる¹⁰．

式 (5) に基づいて参照和音 X を更新することで，最適な X の事後分布を求めることができる．このモデルを，ベイズ更新 (BU; Bayesian updating) モデルと呼ぶ．BU モデルは 12 種類の参照和音の分布をもち，観測した和音に基づいて分布を逐次更新するモデルである．

ベイズスイッチモデル: Bayesian updating and switching (BS) model BU モデルにおける参照和音 X の事後確率のもつ性質に着目して，さらなる拡張を試みた．観測した和音が予想外だったとき，事後確率は低い値を示す．事後確率が低くなったときにリセットする閾値 θ を導入することで，参照和音をより適応的に更新することができる．

$$P(X|c_{1:t}) \propto \begin{cases} P(c_t|X, c_{t-1})P(X|c_{t-1}) & (t = 2 \vee P(X|c_{1:t}) < \theta), \\ P(c_t|X, c_{t-1})P(X|c_{w:t-1}) & (\text{otherwise}) \end{cases} \quad (10)$$

ここで w は直前に X のスイッチが生じた時刻を表す．スイッチが起きていないときは $w = 1$ となる．このモデルをベイズスイッチモデル (BS; Bayesian switching model) と呼ぶ．BU モデルを拡張したため最適な X を求めることはできないが，観測から見積った X の曖昧さに素早く適応することができる．

二次マルコフモデル: Second-order Markov (2M) model もう一つのマルコフモデルの拡張方法として，マルコフ連鎖の次数の拡張が挙げられる．多次元尺度構成法や階層クラスタリングの結果では，第 2 和音と第 3 和音間の音程による効果と，第 3 和音と第 4 和音の音程による効果が類似していた (図 10 を参照)．このパターンは 1 時刻前だけでなく，2 時刻前の和音による効果を拡張した 2 次マルコフモデル (2M モデル) によっても表現できる． $t \geq 3$ のとき，式 (3) は以

¹⁰ 本来は 12 種類の和音それぞれに対してパラメータを割り振り影響を確認するほうが望ましいが，データ数の問題上推定精度が落ちることから避けることとした．また，図 9 を参考にして 5 度の和音に対してもパラメータを付与することは可能であるが，シンプルな定式化を優先するため本研究では扱っていない．

下のようになる．

$$E(c_t = \mathbf{c}) = \sum_X P(c_t = \mathbf{c} | X, c_{t-1}, c_{t-2}) P(X) \quad (11)$$

ここで，常に $X \equiv c_t$ を満たすものとする．すなわち，2M モデルの遷移行列 \mathbf{T} も $12 \times 12 = 144$ 通りの組み合わせを持つ． \mathbf{T} は $i - 1$ 半音低い和音 \tilde{c}_{t-2} を時刻 $t - 2$ に観測し， $j - 1$ 半音低い和音 \tilde{c}_{t-1} を時刻 $t - 1$ に観測したときの期待の大きさ $\mathbf{T}(i, j)$ を要素としてもつ行列となる．

$$\mathbf{T} = \begin{bmatrix} P(\tilde{c}_t=0|\tilde{c}_{t-1}=0, \tilde{c}_{t-2}=0) & P(\tilde{c}_t=0|\tilde{c}_{t-1}=1, \tilde{c}_{t-2}=0) & \dots & P(\tilde{c}_t=0|\tilde{c}_{t-1}=11, \tilde{c}_{t-2}=0) \\ P(\tilde{c}_t=0|\tilde{c}_{t-1}=0, \tilde{c}_{t-2}=1) & P(\tilde{c}_t=0|\tilde{c}_{t-1}=1, \tilde{c}_{t-2}=1) & \dots & P(\tilde{c}_t=1|\tilde{c}_{t-1}=11, \tilde{c}_{t-2}=1) \\ \vdots & \vdots & \ddots & \vdots \\ P(\tilde{c}_t=0|\tilde{c}_{t-1}=0, \tilde{c}_{t-2}=11) & P(\tilde{c}_t=0|\tilde{c}_{t-1}=1, \tilde{c}_{t-2}=11) & \dots & P(\tilde{c}_t=0|\tilde{c}_{t-1}=11, \tilde{c}_{t-2}=11) \end{bmatrix} \quad (12)$$

ただし $t = 2$ のときは事前知識を持たないものとし，全ての \tilde{c}_{t-2} について平均を取った値を採用した．

その他のモデル 行動分析の結果を満たさない他のモデルも，比較対象とした．

最もシンプルなモデルは標的和音の音高によって DOR が決まるモデルである (Pitch-based decision model; 1P model) ．

$$E(c_t = \mathbf{c}) = P(c_t = \mathbf{c}) \quad (13)$$

標的和音に加え，その1時刻前の和音の音高によって決まるモデル (pitch-based Markov model; 2P model) も定義できる． $t \geq 2$ のとき，式 (3) は以下を満たす．

$$E(c_t = \mathbf{c}) = P(c_t = \mathbf{c} | c_{t-1}) \quad (14)$$

一次マルコフモデル (the first-order Markov model; 1M model) も候補とした．1M モデルは，2P モデルにおける期待行列の斜め方向の成分が全て同じ値を取るようにしたモデルに相当する．

4.2.2 モデルの検証

実験で得られた参加者の行動データとモデルの関係を評価するために、行動データに基づいてモデル内部のパラメータを推定し、当てはまりの良さを比較した。

パラメータの推定 連鎖和音の最終音 c_τ を聴いたときの主観的適切度 (DOR) の回答は以下の9通りである。

$$A(c_\tau) \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

参加者の行動選択 A とモデルの出力である適切度の大きさ E は以下の単純な線形関係にあると仮定した。

$$A(c_\tau) = 10 \cdot E(c_\tau) \quad (15)$$

この仮定は適切度の大きさが $0 \leq E(c_\tau) \leq 1$ の範囲内にあり、適切か不適切かどうかとも判断できない中立の状態 $A(c_\tau) = 5$ において $E(c_\tau) = 0.5$ を満たす最も単純な仮定である。本研究ではモデル間の比較やパラメータの傾向の分析が目的であり、 A と E の対応関係が単調増加関数の関係であれば議論への影響はない。そのため、この単純な仮定を採用することとした¹¹。

モデルのパラメータは参加者の行動選択 A とモデルの出力 E の相互エントロピー $L_{(A,E)}$ が最小となるように求めた。

$$(\hat{\mathbf{T}}, \hat{\kappa}, \hat{\theta}) \approx \arg \min_{\mathbf{T}, \kappa, \theta} L_{(A,E)} \quad (16)$$

$$L_{(A,E)} \equiv - \sum_n \sum_{c_\tau^n} A(c_\tau^n) \log E(c_\tau^n | \mathbf{T}, \kappa, \theta) \quad (17)$$

ここで n は試行のインデックスとする ($n = 1, 2, 3, \dots, N$)。

パラメータを以下のように段階的に更新することで推定した。

1. ランダムなパラメータの初期値を生成する

¹¹本研究ではモデルの単純さを優先したが、例えばボタン押しの個人差のモデルとしてシグモイド状の関数を利用することも考えられる。

2. ランダムな順番でパラメータのひとつを選択し、値を増減させた場合の相互エントロピー $L_{(A,E)}$ を計算する¹²
3. 相互エントロピーがより小さくなるパラメータを採用する
4. 全てのパラメータを順に 2~3 を繰り返す
5. 相互エントロピーが一周回って更新されなかったら、更新を止める
6. 上記までの手順を 100 回繰り返す
7. 得られたなかで最小値を示すパラメータを選ぶ

各パラメータの最小更新幅は \mathbf{T} , κ , θ それぞれ 0.1, 1/12, 0.05 とした。ただし、同一の和音が連続する刺激に対して評価を求めた試行とそれに対応する期待行列の対角成分は、実験参加者が回答に混乱したという感想を報告してきたため、本分析の対象から除外した（付録 B.2 参照）。

モデルの評価 パラメータ推定で求められた相互エントロピー $L_{(A,E)}$ はどれだけモデルが参加者のデータを近似できたかを示す。しかし、複雑さの異なるモデル同志について相互エントロピーを用いて直接比較することはできない。AIC[64]等の尤度に対するパラメータの数を考慮した補正方法は、評価尺度と内的な確率との間に単純な線形の仮定を置いたため、慎重になる必要がある。

実験参加者個人ごとの分析ではデータ数が限られているため、2連鎖条件及び3連鎖条件のデータを訓練データとし、外挿条件に当たる4連鎖条件のデータを試験データとしてモデル間の比較を行った（図 12）。

一方全参加者のデータを用いた分析では10分割交差検定を行った。データをランダムに10個のサブセットに分割し、9個のサブセットでパラメータ推定を行って残りで試験を行うことを10通り全ての組み合わせについて実施し評価した（図 13）。

¹²BS モデルが特殊な拡張を含んでおり微分不可能であるため、本研究ではこのような手順を採用した。

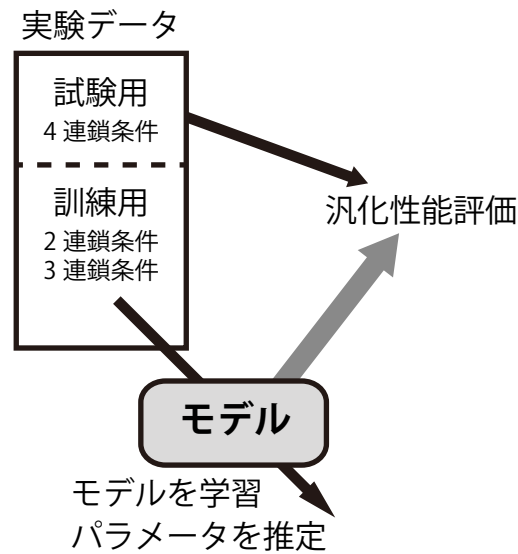


図 12 個人ごとのモデル評価のフロー

数値実験による精度検証 BU モデル及び BS モデルにおいて，参照和音 X は期待行列の混合比を調節する確率変数として機能する．このような直接観測できない隠れ変数を持つモデルでは，パラメータ推定の結果が局所解に陥る可能性がある．そこで，パラメータ推定の精度について確認するため，数値実験を行った．ここでは，データ数の最も少ない個人分析を念頭において精度を確認した．

まずランダムなシミュレーションパラメータをモデルに当てはめて，実験と同じ観測データ（同じ和音の連続を除いた 2 連鎖条件及び 3 連鎖条件の計 253 通り）が与えられた場合の人工的な行動データを作成した．その後，参加者ごとのモデルパラメータの推定と同様に，人工的な行動データから逆に元のパラメータを推定した（図 14 参照）．これらの工程をさらに 100 回繰り返し，真のシミュレーションパラメータと推定したパラメータの平均誤差と標準偏差を算出した¹³．

¹³ここでは，人工データの生成過程においてノイズを加えておらず，本数値実験の結果は与えられた行動実験のデータに対する最尤解からの推定誤差を評価していることになる．

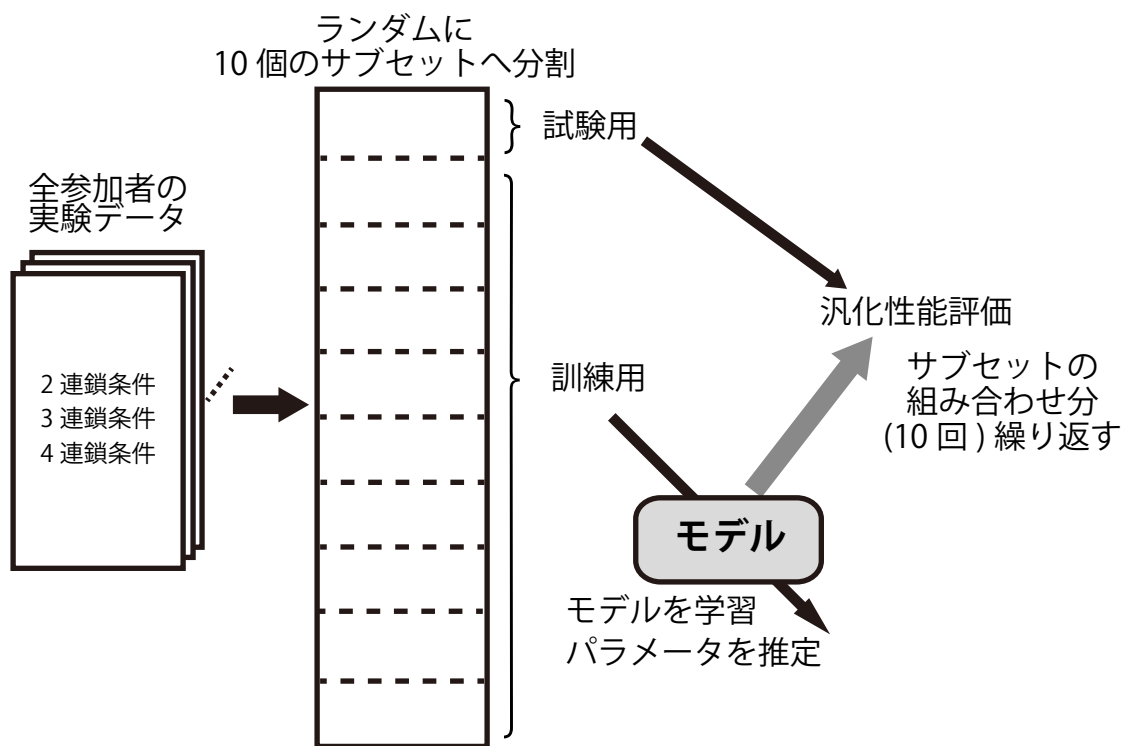


図 13 交差検定によるモデル評価のフロー

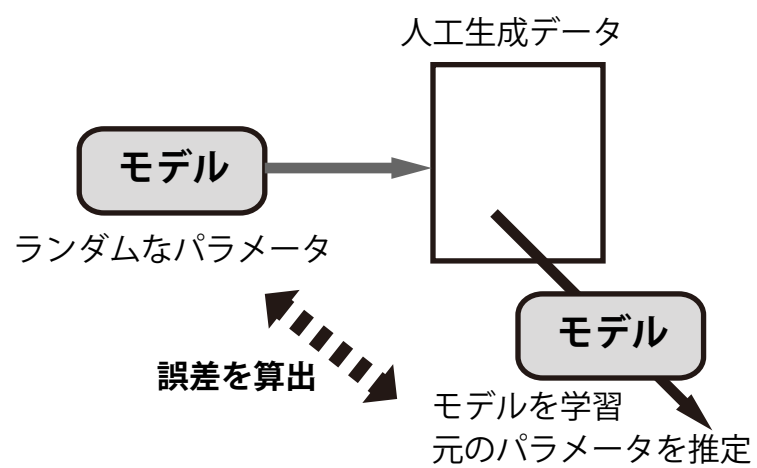


図 14 モデル推定精度検証のための数値実験のフロー

4.3 結果

4.3.1 パラメータ推定の精度

ランダムなパラメータを与えた場合に，BU モデル及び BS モデルのパタメータ推定において生じる誤差を見積もった数値実験の結果を，図 15 に示す．実験の結果，両モデルとも推定結果に大きなバイアスは見られなかった．しかし，BS モデルのほうが BU モデルと比べてやや誤差が大きかった．

4.3.2 個人ごとの分析

個人ごとに 2-chords 条件及び 3-chords 条件の DOR でモデルを訓練し，4-chords 条件の DOR に対する汎化性能を確かめた．訓練データ，試験データそれぞれに対する相互エントロピーを図 16 に，試験データに対する値について統計的検定を行った結果を表 2 に示す．試験データに対し，BU モデル及び BS モデルが 2M モデルと比べて相互エントロピーの値が低くなり（図 16(b)），その差は有意であった (sign test, $p < 0.05$)．しかし，第 3 章の行動分析の結果と矛盾する 1M, 2P, 1P モデルの相互エントロピーも比較的低い値となった．また，BU モデルと BS モデルの間に有意差は認められなかった．このときの BU モデル，BS モデルの推定されたパラメータをそれぞれ図 17 と図 18 に示す．

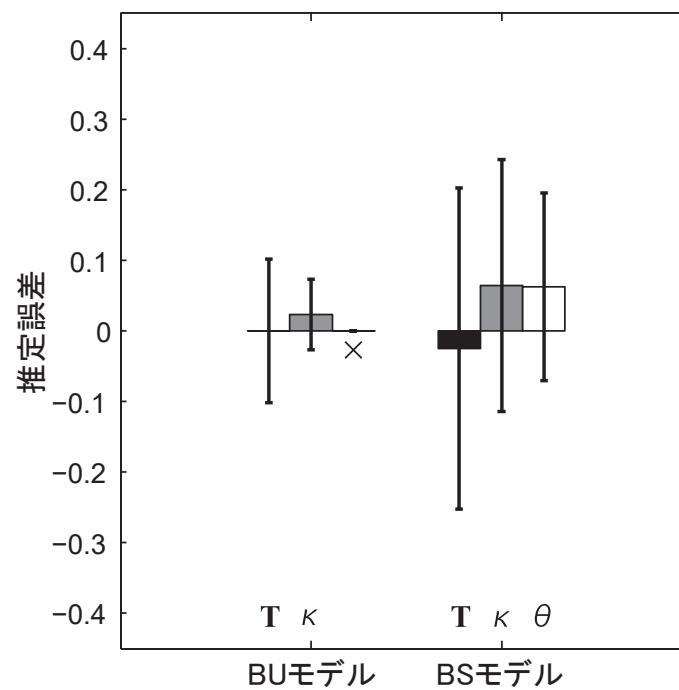


図 15 パラメータの推定誤差

モデルの対応するパラメータの真値からの平均バイアスと標準偏差を表す。×印は対応するパラメータがないことを表す。なお、隠れ変数である参照和音を含まないその他のモデルでは、推定誤差は生じない。

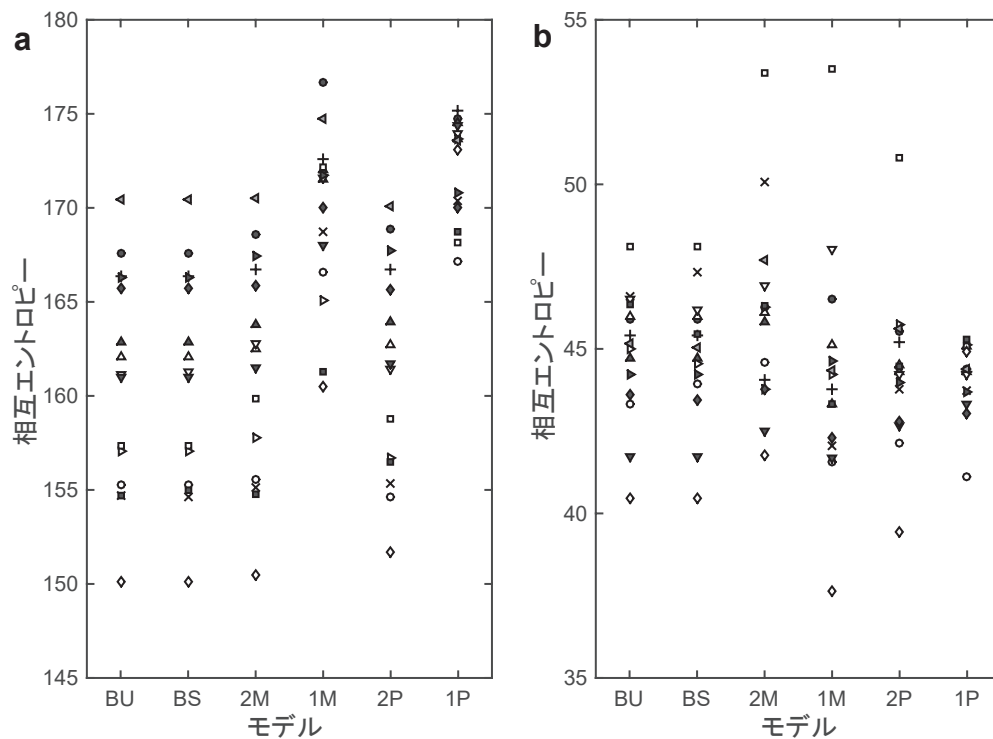


図 16 個人ごとに推定した場合のモデル間の性能比較

(a) 各モデルの訓練データ (2 連鎖条件及び 3 連鎖条件) に対する相互エントロピー．値が低いほどよくデータに対しフィッティングできていることを意味する． (b) 試験データ (4 連鎖条件) に対する相互エントロピー．値が低いほどデータをよく近似できている (汎化性能が高い) ことを意味する．同一参加者に対するデータは同じマーカーでプロットした．

表 2 個人分析の試験データに対する相互エントロピーの統計的検定

モデル 1	モデル 2	sign test			
		1 > 2	1 < 2	N	p -value
BU	BS	5	2	7	0.453
BU	2M	3	12	15	0.035*
BU	1M	11	4	15	0.119
BU	2P	11	4	15	0.119
BU	1P	11	4	15	0.119
BS	2M	2	13	15	0.007**
BS	1M	11	4	15	0.119
BS	2P	11	4	15	0.119
BS	1P	11	4	15	0.119
2M	1M	11	4	15	0.119
2M	2P	12	3	15	0.035*
2M	1P	12	3	15	0.035*
1M	2P	5	10	15	0.302
1M	1P	6	9	15	0.607
2P	1P	9	6	15	0.607

(* for $p < 0.05$, and ** for $p < 0.01$)

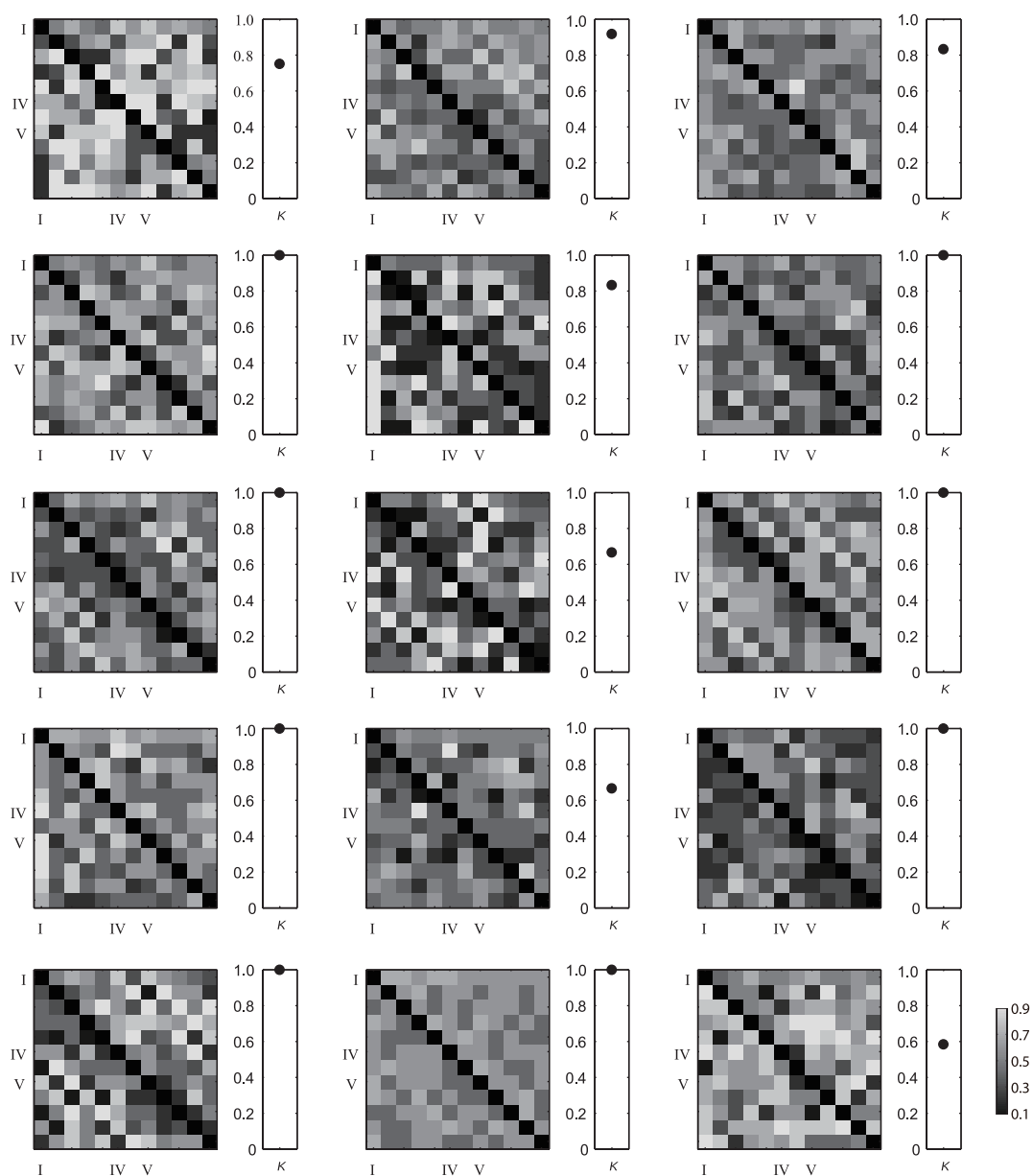


図 17 個人分析における BU モデルの推定パラメータ

推定された期待行列 $\hat{\mathbf{T}}$ 、及び第 1 和音への重み付けパラメータ κ を示す。期待行列の各パラメータの大きさはグレイスケールで表す。なお、期待行列の縦軸横軸は共に参照和音からの音程とし、黒で塗りつぶされた対角成分は推定対象から除外されている。

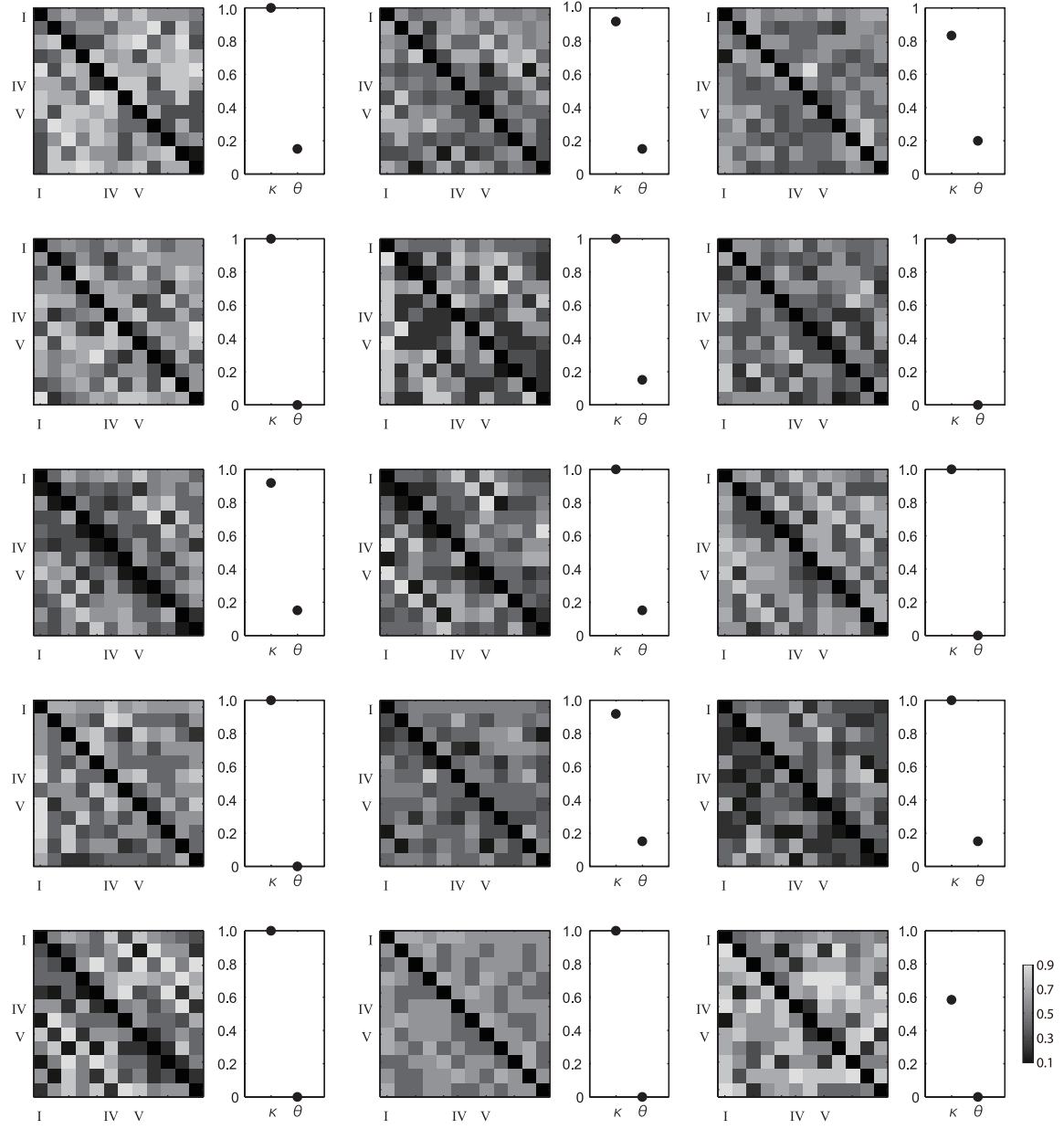


図 18 個人分析における BS モデルの推定パラメータ
推定された期待行列 $\hat{\mathbf{T}}$, 第 1 和音への重み付けパラメータ κ , 及びスイッチングパラメータ θ を示す。期待行列はグレイスケールで表す。

4.3.3 交差検定によるモデルの汎化性能比較

全参加者，全条件のデータを用いて，10 分割交差検定を行った結果を図 19，統計的検定の結果は表 3 に示す¹⁴．個人分析の結果と同様に，試験データに対する相互エントロピーについて，BU モデル及び BS モデルとその他のモデルの間に有意差が見られた (sign test, $p < 0.05$; two-tailed paired t-test, $p < 0.05$ ¹⁵)．BU モデルと BS モデルの間には有意差はなかった．推定された両モデルのパラメータを図 20 に示す．また，BU モデルによる DOR の予測値と実験参加者の平均 DOR の散布図を図 21 に示す．両者の相関係数は $r = 0.97$ であった．

¹⁴本研究では参加者ごとで得られるデータ数が限られているため，交差検定は全参加者のデータを用いて行った．

¹⁵ここではノンパラメトリック検定である符号検定と，パラメトリック検定である T 検定の両方を実施した．交差検定では理想的にはデータの散らばりに対し正規性を仮定できるが，データサブセットのよってばらつきが見られたため符号検定も併記する

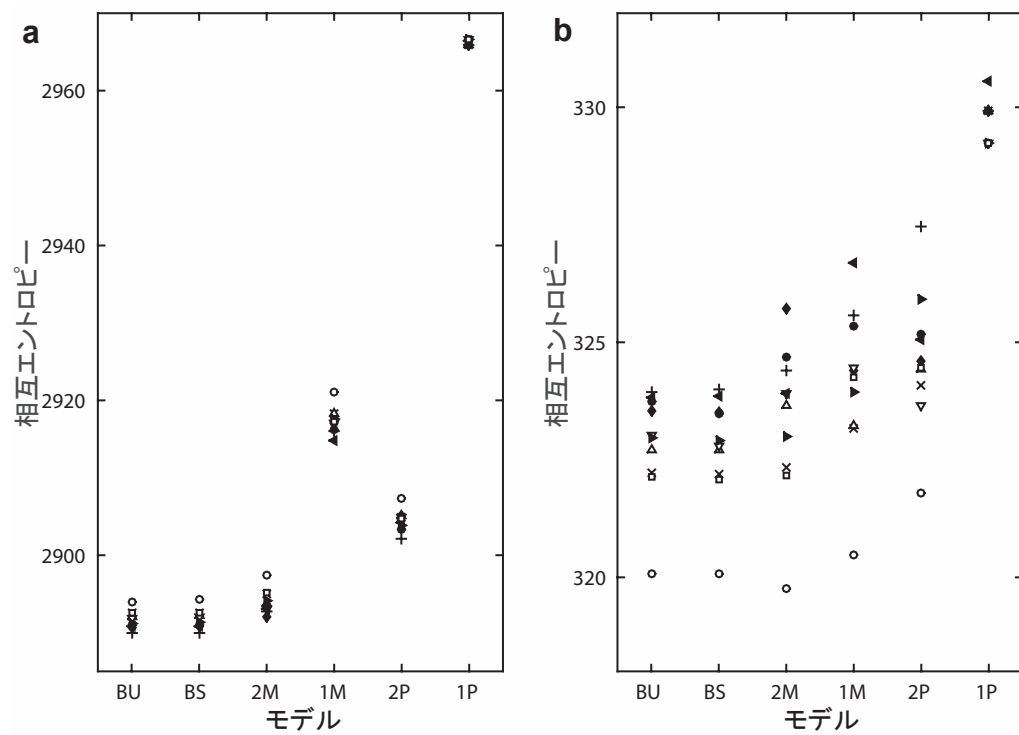


図 19 10 分割交差検定によるモデル間の性能比較

(a) 訓練データに対する相互エントロピーと (b) 試験データに対する相互エントロピー。同一のデータサブセットに対する値は同じマーカーで表す。

表 3 交差検定の試験データサブセットに対する相互エントロピーの統計的検定

モデル 1	モデル 2	sign test				paired T-test		
		1 > 2	1 < 2	N	p -value	T	N	p -value
BU	BS	6	4	10	0.754	1.570	10	$p = .151$
BU	2M	1	9	10	0.022*	-2.343	10	$p = .044^*$
BU	1M	0	10	10	0.002**	-5.595	10	$p < .001^{***}$
BU	2P	0	10	10	0.002**	-6.628	10	$p < .001^{***}$
BU	1P	0	10	10	0.002**	-22.433	10	$p < .001^{***}$
BS	2M	1	9	10	0.022*	-2.447	10	$p = .037^*$
BS	1M	0	10	10	0.002**	-5.702	10	$p < .001^{***}$
BS	2P	0	10	10	0.002**	-7.215	10	$p < .001^{***}$
BS	1P	0	10	10	0.002**	-23.511	10	$p < .001^{***}$
2M	1M	2	8	10	0.109	-2.201	10	$p = .055$
2M	2P	2	8	10	0.109	-3.029	10	$p = .014^*$
2M	1P	0	10	10	0.002**	-13.661	10	$p < .001^{***}$
1M	2P	3	7	10	0.344	-1.366	10	$p = .205$
1M	1P	0	10	10	0.002**	-12.037	10	$p < .001^{***}$
2P	1P	0	10	10	0.002**	-11.680	10	$p < .001^{***}$

(* for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$)

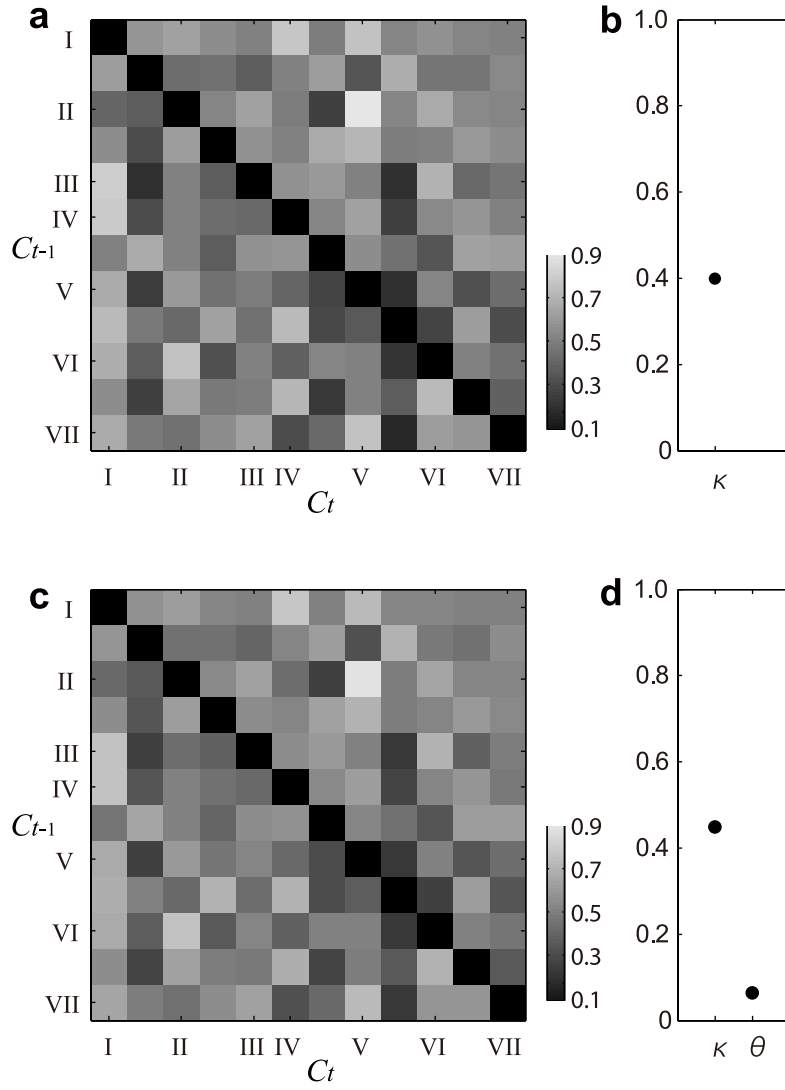


図 20 交差検定により推定されたパラメータ

上段 (a-b) に BU モデル，下段 (c-d) に BS モデルの推定されたパラメータを示す． (a) と (c) は期待行列 $\hat{\mathbf{T}}$ の推定値， (b) と (d) はその他のパラメータの推定値である．ただし BU モデルではスイッチングパラメータ θ は存在せず常に 0 と等しい．

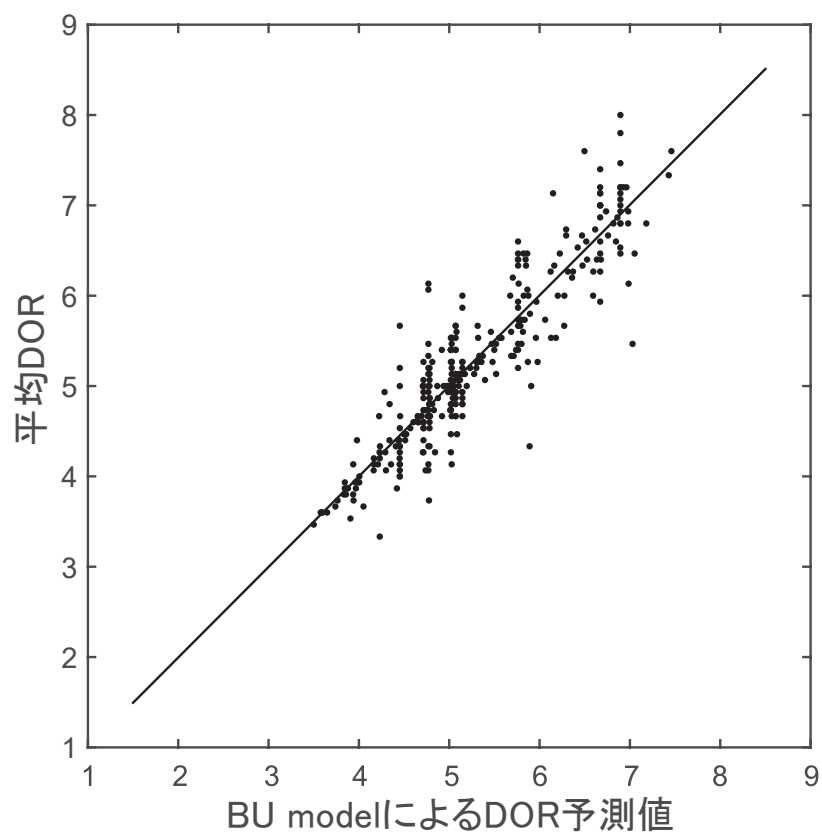


図 21 BU モデルによる DOR 予測値と参加者間平均 DOR の散布図
交差検定によって推定された BU モデルのパラメータを用いて DOR の予測値を算出し、
参加者の平均 DOR との関係プロットした。

4.4 追加分析による検証

4.4.1 二値強制選択課題と見なした分析

評価尺度を用いたパラメータ推定は、実験参加者の内的な期待の大きさの個人差のほかに、ボタン押しの傾向の違いも含まれる。そこで評価尺度を二値強制選択と見なし、全参加者のデータを用いて同様にモデルを推定した。実験ではDORの真ん中にあたる「5」を「適切、不適切のどちらでもない」ときに押すように教示しており、5を境に高いときと低いときをそれぞれ1, 0の二値に変換した。なお、DOR上の「5」の行動選択は、二値強制選択課題の場合理想的に1と0を等確率で取ることになる。このため、パラメータの推定やモデル間の性能比較において全てのモデルに対して情報を与えないため、本分析では除外して扱った。モデルの性能はAICにより評価した。

結果を図22に示す。他の分析と同様にBUモデルとBSモデルが他モデルよりも低い値を示した。また本比較においては、BUモデルが一番低い値を示した。この分析で推定されたパラメータを、図23に示す。

4.4.2 キープロファイルの導出

BUモデル及びBSモデルにおける参照和音は、連鎖和音の聴取によって生じる内的な音程の基準である。すなわち、音楽理論における調(の主和音)に相当すると解釈できる。

先行研究では調によって生じる特徴的な指標として、プローブ音法によって導かれるキープロファイル[10]が用いられてきた。キープロファイルはヒトの内的な調性的階層を示す手法としてだけでなく、調の自動推定モデルにも参考とされてきた[65, 6, 66]。そこでモデルから理論的なキープロファイルを導出し、先行研究で得られてきたキープロファイルとの比較を行った。なお、BUモデルとBSモデルでほとんど推定されたパラメータや汎化性能に差が見られなかったため、より単純なBUモデルを用いて検証した。

BUモデルにおいて、期待行列 \mathbf{T} (式8)はある調が与えられたときの和音進行に対する主観的適切度を意味する。そのため、時刻 t における和音 j に対する

主観的適切度 $E_{c_t=j}$ は \mathbf{T} から以下の計算式で見積もることができる¹⁶.

$$E_{c_t=j} = \frac{1}{11} \sum_i \mathbf{T}(i, j) \quad (18)$$

図 24(a) に BU モデルから推定されたキープロファイルを示す．先行研究で得られてきたキープロファイル（図 24(b) 及び図 24(c)）と同様に主要 3 和音に対し高い評価値を示した．その一方で，他の長音階を構成するピッチクラスを根音とする和音についてもやや高くなる傾向が見られた．

¹⁶式 18 の正規化のための定数がピッチクラスの数 12 でなく 11 であるのは，自己遷移に相応する対角成分を含まないためである．

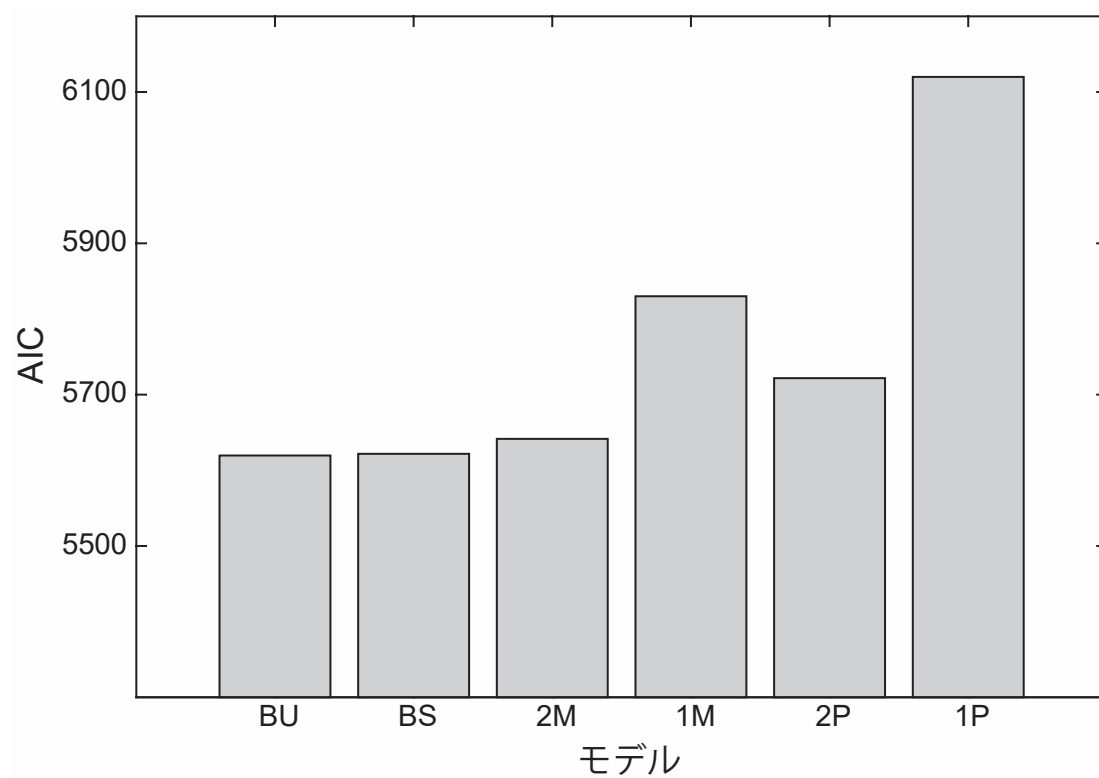


図 22 二値強制選択課題と見なした場合のモデル間性能比較
各モデルについて、全データを用いて算出した AIC の値を示す。値が低いほど、モデルの当てはまりがよい。

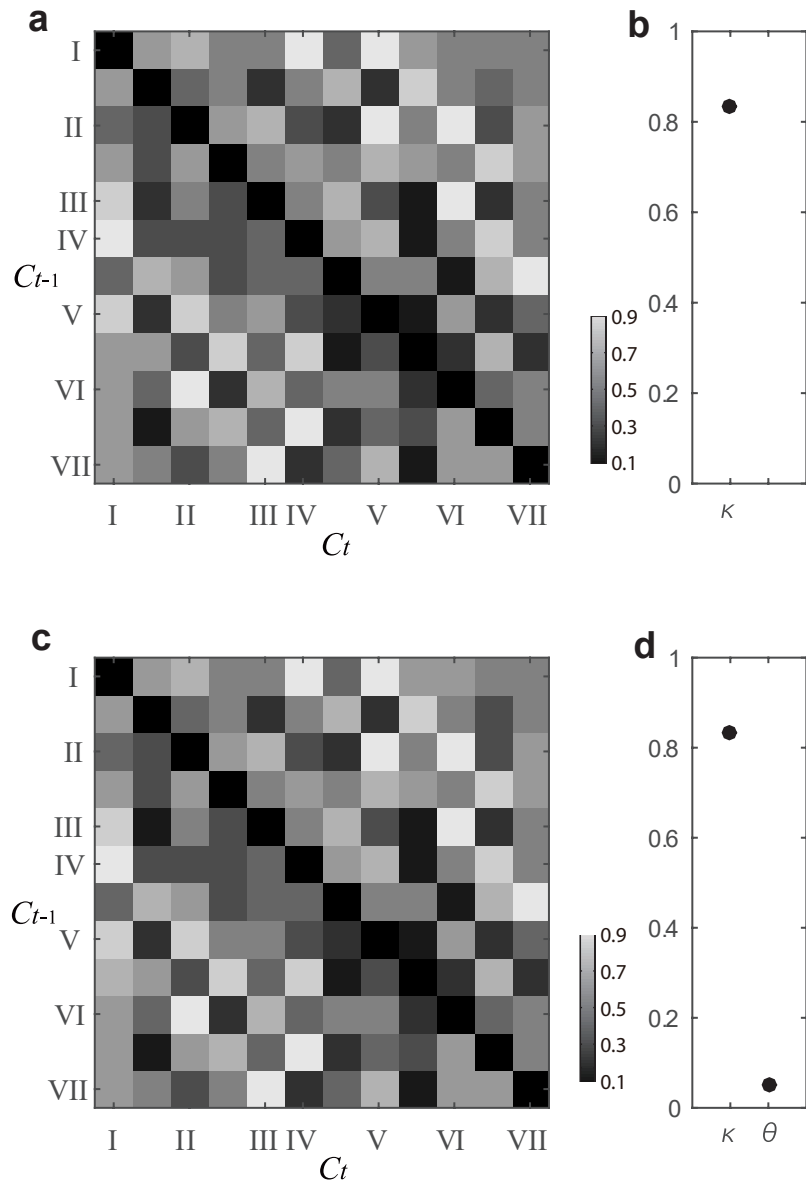


図 23 二値強制選択課題と見なした分析により推定されたパラメータ
 図 20 同様、上段 (a-b) に BU モデル、下段 (c-d) に BS モデルの推定されたパラメータを示す． (a) と (c) は期待行列 $\hat{\mathbf{T}}$ の推定値， (b) と (d) はその他のパラメータの推定値である．

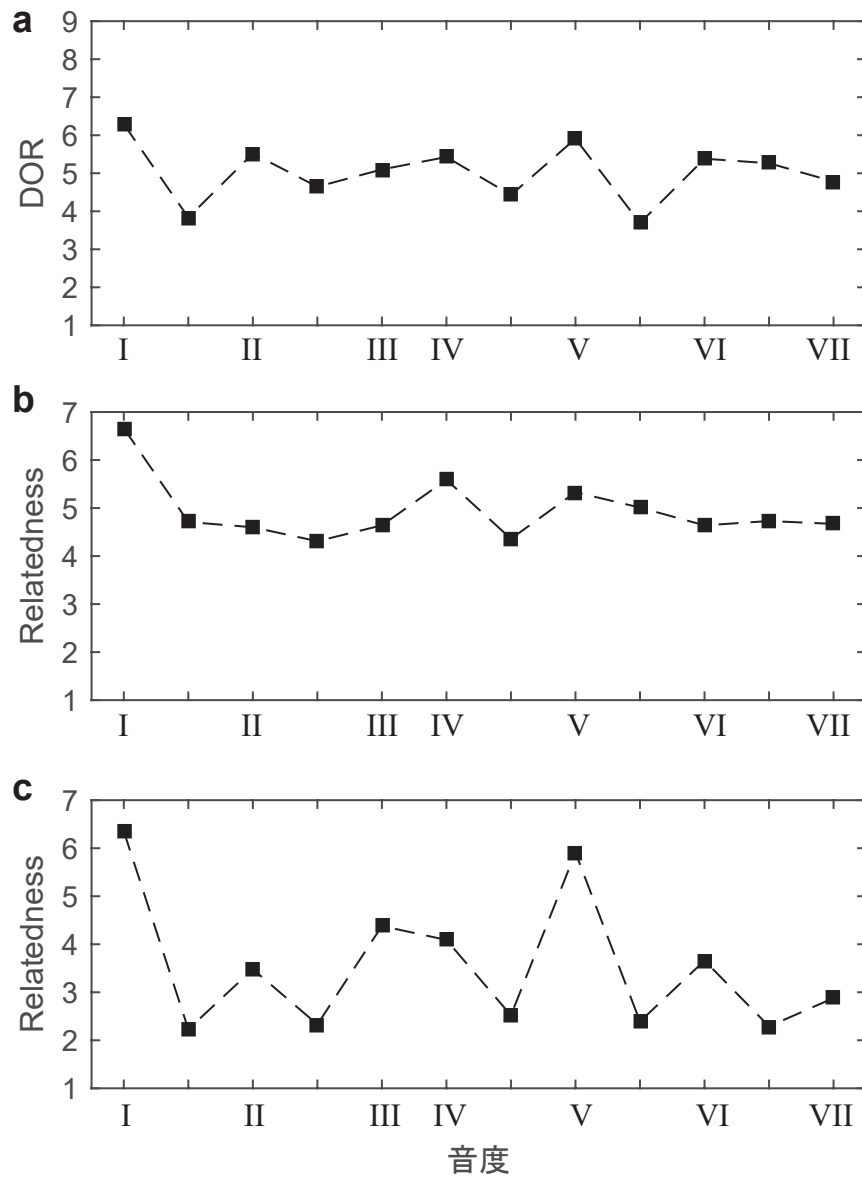


図 24 キープロファイルの比較

(a) BU モデルの期待行列から算出したキープロファイル. (b) 長調の文脈における長三和音に対するキープロファイル (Krumhansl(1990)[1](p. 171) より引用). (c) 長調の文脈における単音に対するキープロファイル (Krumhansl & Kessler (1982)[10] (p. 343, Figure 2) より引用). なお, (b) 及び (c) では 7 段階の評価尺度が用いられている.

4.5 考察

本章では連鎖和音聴取時に和音進行に対して生じる主観的な適切度 (DOR) について、計算論的アプローチにより内的な生成メカニズムをモデル化した。モデルの候補を DOR を用いて評価したところ、BU モデル及び BS モデルが支持される結果が得られた。これら提案モデルは、前章の行動分析から得られた DOR が音高ではなく音程に影響を受けること、直前の和音からの音程に依存した類似性を持つこと、直接隣接しない音程からも影響を受けることの3つの示唆を満たすモデルであった。推定されたモデルから、参加者が内的な音程の基準 (参照和音) を保持し、その基準と和音進行に応じて期待を生じている可能性が示唆された。また、音程の基準自体も観測した連鎖和音に応じて適応的に更新していると考えられる。

提案モデルが保持する期待行列 (式 (5) 参照) は、ある参照和音を仮定したときに呈示された和音進行がどれだけ尤もらしいかを表す¹⁷。全参加者のデータを用いて推定した BU モデルの期待行列について、表 1 で示した音楽理論上の接続頻度に基づいて分類したプロットを図 25 に示す。本研究では長和音しか用いていないため、この分類には厳密には当てはまらない進行が多いが、最も典型的なグループに属する和音進行では期待が高い傾向が確認できる。一方で根音が音階外となる和音を含むその他の分類についても、期待の高い進行が含まれていることが伺える。表 4 に期待行列の確率が上位及び下位の進行を例示する。いくつかの音階外の和音進行 ($V\sharp$ から IV, $VI\sharp$ から VI など) において、ドミナント進行 (V から I) よりも高い期待を示す傾向が見られた。これらの結果は最大 4 連鎖までの短い連鎖和音に対する参加者の行動データから最尤推定された値である。そのため数値を一般化した議論は困難であるが、計算論的アプローチによって従来の音楽理論では記述されていなかった音楽的期待の側面が明らかにできる可能性を示している。すなわち、知覚実験から推定された計算モデルに基づいて、音楽理論を拡張することができると考えられる。

また推定したモデル内の期待行列から求めたキープロファイルは、プローブ音

¹⁷音楽情報科学の研究では、和音進行をマルコフ過程の遷移確率によってモデル化してきたが [37, 38], この期待行列では各要素が期待の大きさを表している。

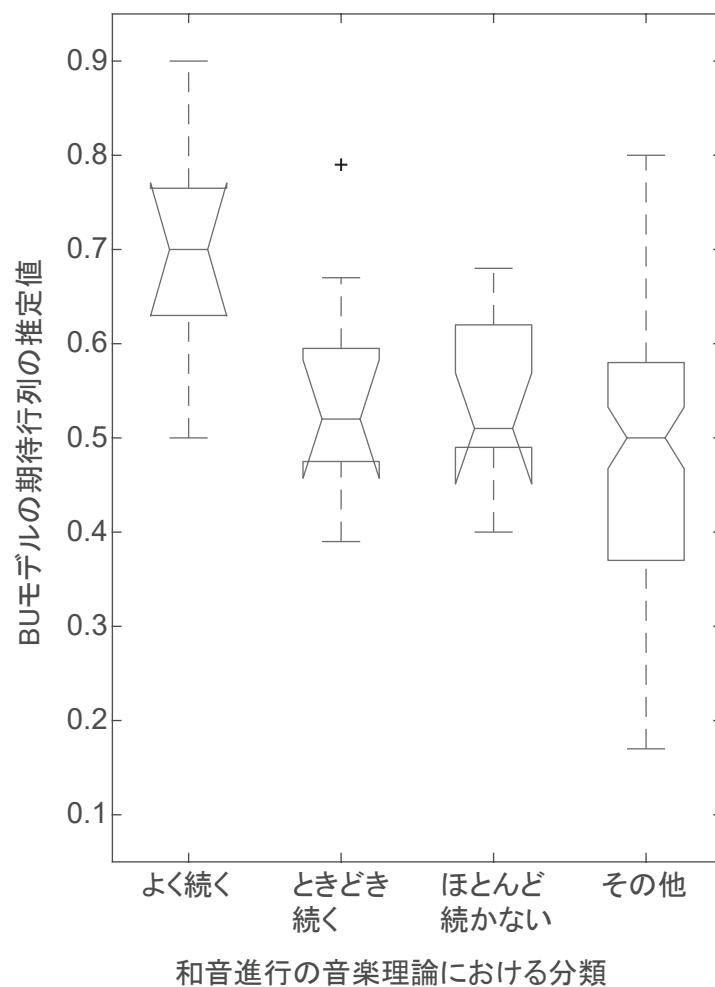


図 25 期待行列と音楽理論の比較

推定した BU モデルの期待行列 (図 20) の値を，図 1 に記述される 3 種類の和音進行及びそれ以外の計 4 つのカテゴリーに分けてプロットした．

法から得られたものとよく似たパターンを示した（図 24）。上述の期待行列の傾向も合わせると、内的な音程の基準は音楽理論における調に相当すると解釈できる。期待のパターンはそれまでの音楽経験に基づいて無意識化に学習し構築されるとされており [14, 16]，西洋音楽内に構造的に含まれる和音進行の頻度と実験的に得られた期待のパターンの一致が指摘されてきた [1, 67]。本研究の結果は、この仮説を支持するものだと言える。

一方，個人ごとに推定された期待行列間にはばらつきが見られた（図 17 及び図 18 参照）。全参加者の期待行列の平均値と標準偏差を図 26 に示す。本実験では，9 段階の評価尺度を用いており，評価尺度と内的な確率との対応関係が（中点を除いて）参加者ごとに異なる可能性があるため，厳密な個人差の評価は難しい。しかし図 26(b,d) に見られるように，参加者間の標準偏差は期待行列の位置

表 4 期待行列の上位及び下位の和音進行

推定した BU モデルの期待行列（図 20）における上位及び下位それぞれ 10 位までの確率の和音進行を表す。

上位	1	2	3	4	5	6	7	8	9	10
進行	II-V	III-I	IV-I	I-IV	VII-V	VI-II	I-V	V \sharp -IV	VI \sharp -VI	V \sharp -I
確率	0.9	0.8	0.79	0.78	0.76	0.76	0.75	0.73	0.73	0.72
下位	1	2	3	4	5	6	7	8	9	10
進行	VII-V \sharp	III-II \flat	III-V \sharp	V-V \sharp	VI-V \sharp	VI \sharp -IV \sharp	V-II \flat	IV-V \sharp	II-IV \sharp	VI-II \flat
確率	0.17	0.2	0.2	0.2	0.21	0.22	0.24	0.25	0.25	0.26

（太字は表 1 に示される典型的な和音進行を表す）

によって異っており、個人差の存在を示唆している¹⁸¹⁹。これまで、文化的背景の異なる実験参加者間において音楽的期待のパターンや程度に差があることが報告されている [68, 69] 一方、基本的な傾向自体には教育的な訓練の有無による差は認められてない [70]。そのため本結果は、教育的訓練の経験よりも参加者の好みや聴取経験の差に基づく可能性がある（付録 B.1 参照）。

提案モデルにおける内的な音程の基準の和音進行に対する期待への寄与は、活性化拡散モデル [14] における調と和音の関係とも矛盾しない。活性化拡散モデルに従えば、入力された和音は関係の深い調を重み付けする形で活性化し、続いて活性化した調が関係の深い和音を活性化して期待を大きくする。このような活性の伝搬が、さらに次時刻の和音の観測によって関係の深い調への重み付けを強化していくことで、内的な調性が構築されていくとされる。提案モデルでも、聴取した和音進行が起こりやすい音程の基準が重み付けされ、その重みが次の時刻の観測によって時々刻々と更新される。

提案モデルでは音程の基準、すなわち内的な調はベイズの定理に基づいて更新される²⁰。ベイズモデルは感覚運動の操作 [71]、視覚 [72]、時間順序の判断 [73]、時間間隔の知覚 [74]、意思決定 [75]、推論 [53] といった、広く様々な知覚・認知現象を説明するモデルとして提案されている。楽曲から調を推定するモデルとしてもベイズモデルが有用であることが示されており [6]、音楽の知覚においても同様な情報処理が行われている可能性がある。調を逸脱する音に対する脳活動である ERAN の活動源と推定される、下前頭回 [2, 22] を含む外側前頭前皮質と呼ばれる領域は、外的な入力に対しその文脈情報に基づいて行動戦略を切り替える役割を担うと考えられている [76]。本研究の提案モデルは、調の情報を連鎖和音の文脈に応じて適応的に更新しており、この仮説と矛盾しない。また同領域は、同

¹⁸特に I の和音からの遷移において参加者間の推定値のばらつきが小さい傾向が見られる。主和音に相当する I の和音からの遷移への期待が、内的に明瞭である可能性はある。しかし、提案モデルは第 1 和音を I の和音と捉えがちであり（後述のパラメータ κ に関する議論を参照）、I の和音からの遷移に対応するサンプル数は他の和音進行と比べて相対的に多くなる。このため、参加者ごとの推定値もより確からしい結果が得られ、標準偏差が小さくなったと考えるほうが自然である。

¹⁹期待行列の平均値（図 26(a,b)）と標準偏差 (c,d) の関係を見ると、特に高い値を示す和音進行について、標準偏差が低くなる傾向が見られる。このことに関する議論は、付録 B.1 を参照。

²⁰期待と音程の基準の関係については、付録 B.3 を参照。

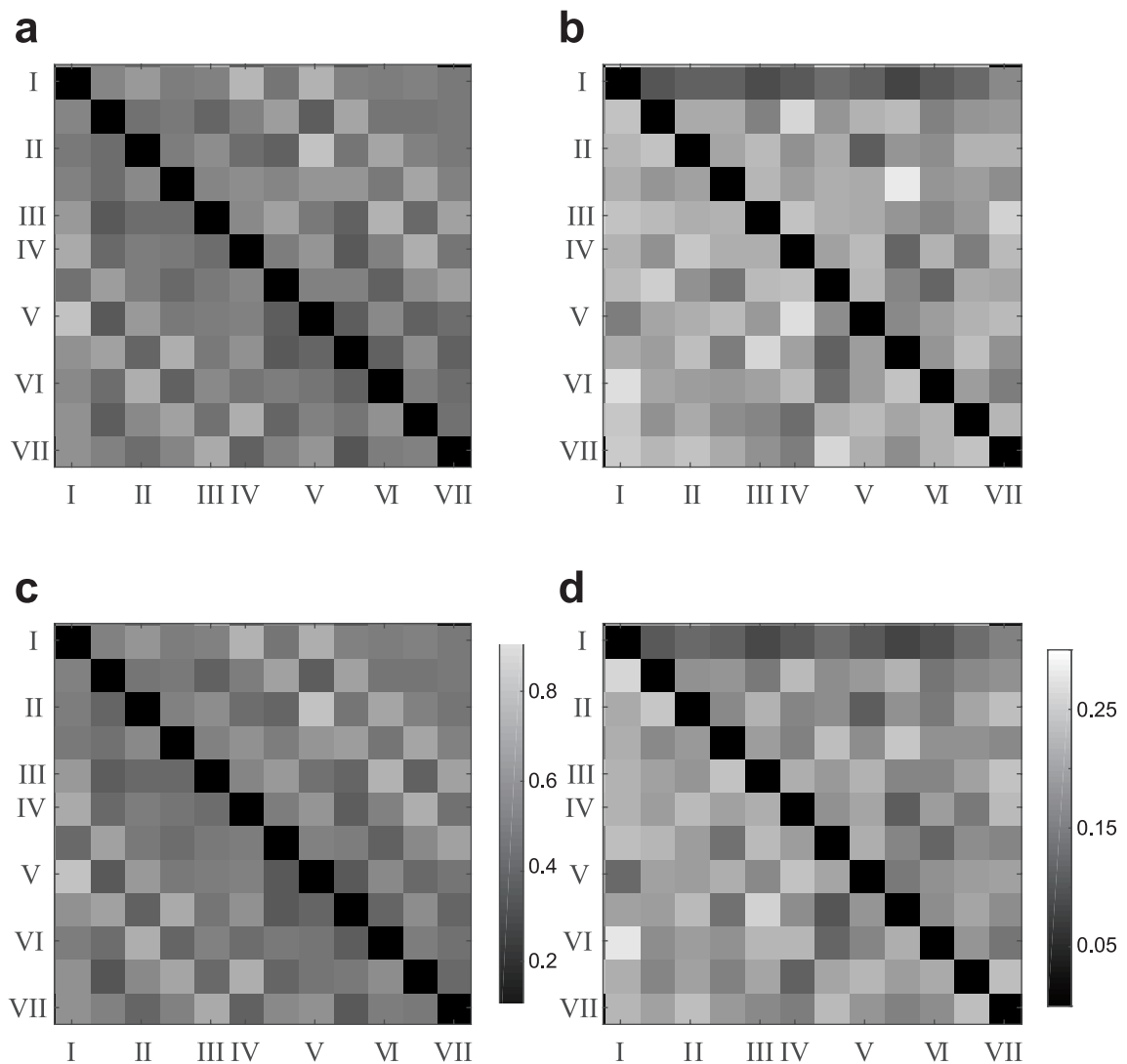


図 26 個人ごとに推定した期待行列の平均と標準偏差
 推定された BU モデルの期待行列の (a) 参加者間平均値と標準偏差 (b), 及び BS モデル
 の (c) 平均値と (d) 標準偏差をグレイスケールで表す。

様にベイズ的な情報処理モデルが提案されている時間順序の判断課題 [77] においても賦活が見られることが報告されている²¹。こうしたことから調性知覚に対して、脳内にあるベイズ的な情報処理基盤が関与していることが示唆される。しかし、本研究だけでは先行研究における脳活動と提案モデルの変数の関係を直接的に示せておらず、今後の研究が待たれる²²。

提案モデルに含まれるパラメータ κ は、第 1 和音に対する重み付けの大きさを意味する。我々の推定結果では、全ての実験参加者において、推定された κ が高い値を示した。このことは 2 番目の和音に対する期待を計算するときに、第 1 和音を調の主和音として捉えがちなことを意味する。すなわち、最も短い 2 つの連鎖和音においても調性の仮定が期待の方向付けに関わっていることが示唆された。

本研究では θ を含む BS モデルも検討したが、BU モデルとの間に汎化性能に関して有意差は認められなかった（個人分析：図 16(b)，交差検定：図 19(b)）。DOR を二値化した分析においては、BU モデルのほうが僅かに低い AIC 値を示した（図 22）。推定された θ はいずれの条件においても低い値が見積もられており、本研究ではパラメータとしての実効性を示せていないと言える。BS モデルは観測を進めるにつれて参照和音が一定の閾値よりも不確定になったとき、これまでの観測履歴を捨てて参照和音の分布をリセットすることができる。これは調性音楽における転調 [80] に対応するモデルであると言える。今回の実験に用いた連鎖和音の連鎖数は最大で 4 つしかないため、このような分布の初期化自体が起りにくく、BS モデルにとっては不利な条件であった可能性がある。

²¹下前頭回は言語の文法処理に関わることで知られている。提案モデルの期待行列は言語における文法に対応すると解釈される。ただし、ベイズ的観点から言語認知を捉えた研究は少ない [78, 79]。

²²例えば、脳波実験の各時刻におけるモデルの内部変数の変動を計算し、変数に相関する脳活動を回帰的に探索する手法が考えられる。これまで ERAN は音楽理論上明確に調を逸脱する刺激を標的とした研究での報告が多かったが、本研究の検証では調の仮定を置かないため、回帰分析を念頭に行動実験と同様の網羅的なデータ収集を行う実験デザインを組むことが望ましいかもしれない。

第5章 総合考察とまとめ

本論文では、連鎖和音聴取時に和音進行に対して生じる音楽的期待に焦点を当て、その生起において内的に構築される調性が関与する過程を明らかにした。第3章では心理実験の統計的分析から、和音進行に対する期待が直前の和音との音程に強く影響を受ける一方、直接隣接していない和音との音程も寄与している可能性を示した。第4章では、計算論的アプローチと呼ばれる構成論的な手法を導入し、前章で得られた知見に基づいて実験参加者の内的な計算過程をモデル化した。モデル選択の結果、観測した和音進行に応じて内的に調が形成され、その調に基づいて期待が決まっていることが示唆された。これらの結果は先行研究によって示されてきた、期待の大きさが音楽理論による機能的解釈と適合するという報告と矛盾ない[13, 17, 46]。

本研究は計算論的アプローチを導入することで、従来研究が陥っていた音楽理論による制約を打開したと言える。調性が内的な変数として存在することは、暗黙な形で前提とされてきた。しかし経験則であるため、調の存在を前提とした知覚実験によって調による影響や音楽理論との整合性を示すだけでは、知覚の全体像を明らかにする上で不十分だったと言える。本研究では和音進行に対する主観的な適切度という行動選択を数理的に妥当な形で説明する上で、調という内的な変数の存在を示唆することができた。

従来の統計的分析と本研究の計算論的アプローチの最大の差異は、変数の取り扱いにある。従来研究では研究者が着目する要因についてコントロールした実験を行った上で、その実験条件を変数として統計的検定を行い、条件間の行動の差について論じてきた。このような還元論的な論法は、ごく単純な知覚現象に対しては有効であるが、単純な関係に落とし込めない多変数に由来する現象に対しては、踏み込んだ解釈が難しくなる。これに対し計算論的アプローチでは、内的に存在する変数を仮定し、変数間の関係を記述したモデルについて行動との一致率を評価している。このような構成論的な論法は、現象に対して統合的に説明を試みるため、内的なメカニズムについて論じやすい。

これまで和音進行に対する期待には、音程に基づく局所的な効果と調性によって生じる大局的な効果の二種類の効果が含まれるとされてきた[46, 63]。本研究

の第3章の結果からも、同様の考察が得られる。一方、第4章で推定された提案モデルは、両方の効果を一つのモデルの中で統合的に説明している。すなわち、局所的な効果は調に関する事前知識を与えなかった場合²³、大局的な効果は調性によって重み付けされた場合の期待の出力に相当する。先行研究では還元論的な立場から条件分けを行い、隣り合う和音間の音程による効果と文脈的に生じる効果が異なる性質を示すという結果を導いてきた。それに対し、本研究では和音の入力から主観的な適切度の出力に至る系全体についてモデル化を行っている。提案モデルは平均的な DOR に対して十分な近似精度を得ており（図 21）、両方の効果が異なる機構によって実現されているというよりは、むしろ調性に関する知識条件の違いによって生じたと考えるほうが妥当であろう。

一方で、計算論的アプローチの適用にはいくつかの障壁がある。特にデータ数について現実的な制約があるため、心理実験に基づく計算モデルの組み立てや評価は簡単ではない。本研究では実験条件の偏りによる推定結果のバイアスを避けるため、短い連鎖和音しか扱うことができなかった。またモデル候補の構築も、構成論的であるために制約がなく、どの程度まで複雑なモデルを検討する必要があるか、判断の客観的根拠が乏しい。本研究ではなるべくシンプルで数理的に妥当性のあるモデルを志向し、行動実験の統計的分析の結果を満たす範囲でモデリングを進めるセミ・トップダウンな方策を取った。しかし提案モデル以上の近似性能を持つモデルの存在は否定できない[51]。今後提案モデルの妥当性を示すために、より長い連鎖和音や異なる種類の和音、異なる和音の長さ等の条件下においても検証していく必要がある。このとき、例えば検証する候補モデルを予備的な実験を元に絞り込んだのち、全候補が対等な条件で比較できるような実験デザインを改めて検討するなどの工夫が必要とされる²⁴。

²³現実的には調に関する情報を与えなくても、第1和音に対する重み付けが生じると考えられる。

²⁴音楽知覚における条件設定の難点のひとつは、物理的な統制が必ずしも知覚的な意味合いでの統制に繋がらないことである。本研究では DOR の分布を統制しなければ、DOR の生起メカニズムを正確に論じられない。例えば、モデル候補についてある実験条件における DOR の予測値を計算し、その予測された DOR の分布の偏りを評価して実験条件を修正するという再帰的な過程の導入などの対策が考えられる。一方で再帰的な過程はモデルに即した議論を過度に促進してしまう側面もあるため、DOR 以外の行動指標による評価もより重要となる。

本研究ではモデルの単純化のため、和音をシンボル化して扱った。しかし、例えば和音の連鎖数や持続長といった時間的な要素も、リズムや拍の知覚とも関連するため主観的な適切度に対して影響を与える可能性が高い。また実際の調性音楽では、和音の構成音の積み重ね方や次時刻の和音への接続、和音を構成する音色といった様々な要素が影響を与えると考えられる。本研究では12平均律における長3和音に限定し、従来の音楽理論と比較できるように和音をシンボリックに扱って検証を進めたが²⁵、今後様々な音楽構成要素との関連を検討しなければならない。

音楽知覚の研究には大きく2つの意義がある。1つは芸術を芸術として知覚するメカニズムの解明である。和音進行に対する期待からの逸脱は情動を誘発する側面があるとされ[81, 82]、調性音楽の持つ芸術的な側面と関りがある。²⁶もう1つの意義は、ヒトの時間的な情報処理メカニズムの考察である。特に提案モデルにおける内的な変数としての調の存在やその逐次的な更新は、音楽聴取のために後天的に獲得した脳機能であるとは考えにくく、ヒトの脳に備わる効率的な情報表現やベイズ的な情報処理の一端を示していると言える。音楽刺激は含まれる変数を外的に操作しやすく、知覚応答を分析することで様々な時間的な情報処理について考察することができる。また音楽理論や音楽の構造研究から生まれたモデル[43, 67]も、ヒトの内的な計算の特性を暗に含んでいる可能性があり、研究の重要なヒントとなりうる。

様々な課題が存在するが、計算論的アプローチは直接観測できない内的な変数に基づくメカニズムを論じる上で有効な手法である。音楽知覚の研究において、作曲者の意図や楽曲の構造ではなく、心理実験に基づいて知覚メカニズムを調べることは順当な手段であると言える。本手法は音楽に限らず様々な知覚現象に対して同様な形で適用可能であり、今後様々な知覚メカニズムの解明に役立つことが期待される。

²⁵本研究は12平均律の範囲で音程に関する循環系を構成できるようにすることで、現実的な試行数内での検証を可能とした。12平均律以外の調律への拡張も今後の課題である。

²⁶20世紀以降発展してきた現代音楽では、無調音楽のように非調性的な響きを探求した作品も存在する。理想的な無調であれば本モデルを当てはめると調（音程の基準）の分布は一様分布になる。

謝辞

本稿の査読をお引き受けくださった、本学 情報科学研究科 数理情報学研究室の池田 和司先生、知能コミュニケーション研究室の中村 哲先生、数理情報学研究室の吉本 潤一郎先生、お茶の水女子大学大学院 人間文化創成科学研究科の伊藤 貴之先生に、厚くお礼申し上げます。本論文は、本学博士後期課程在籍時に思案していたアイデアを元に、九州大学大学院 芸術工学府にて実施した研究を中心とし、いくつかの分析を追加した上でまとめたものです。池田先生には、そのような複雑な経緯にも関わらず査査を快諾いただき、審査のためお力添えをいただきました。先生方による審査を通じて、理学及び工学的視点から多数の貴重なご意見を頂戴し、研究をより一層洗練させることができたと思っております。お忙しい中、どうもありがとうございました。

九州大学大学院 芸術工学府の中島 祥好先生及びジェラード B. レメイン先生には、知覚心理学の視点から様々なご意見を頂戴し、雑誌論文の受理までご尽力いただきました。高木 英行先生には主に数理的側面からアドバイスをいただき、評価尺度を用いる本研究の実験デザインに繋がりました。白石 君男先生には脳活動計測関連のゼミの他、普段の研究室生活において様々なサポートをいただきました。また日常的な研究の議論及び研究室生活において、澤井 賢一博士、藤岡 拓也さん、串田 之和さん、久保 翔平さんをはじめとする中島研究室所属（当時）の皆様、白石研究室の皆様には特にお世話になりました。

本研究における「計算論的アプローチ」は、本学修士課程及び博士課程において論理生命学講座（生命システム学講座）在籍時に習得したスキルを下地としたものです。石井 信先生をはじめ多くのスタッフ、所属学生の皆さまにお世話になりました。特に計算論的認知神経科学についてご指導頂いた吉田 和子博士、モデリングについてアドバイスを頂いた平山 淳一郎博士、日常面において手厚くサポートして下さった作村 諭一先生に深く感謝いたします。

研究アイデアの立案に当たっては、若手研究者有志による自主的な勉強会である「音楽と脳」勉強会に大変お世話になりました。アイデアの草案から論文投稿まで、節目において勉強会のメンバーには様々な形でご協力と後押しを頂いてきました。音楽知覚は魅力的なテーマである一方で、複雑な現象を扱うために曖昧

な議論に終始しやすい側面が否めません。教員に限らず研究員や学生まで幅広い視点から率直な意見交換ができる環境を継続させることが、従来の固定観念に囚われない新しいアイデアを育てる上で最も大事なことだと思います。勉強会の運営をサポート頂いた大阪大学大学院 医学系研究科（当時）の木下 博先生，紹介いただいた京都大学大学院 情報学研究科（当時）の寺村 佳子さんに感謝申し上げます。現在は阪大から電気通信大学に開催場所を移しましたが，一員として勉強会の発展に貢献していけたらと思っています。

本学博士課程を単位取得認定退学したのち，学位取得まで多くの方々に様々な形でサポートをいただきました。所属していた本学バイオサイエンス研究科の稲垣 直之先生及び学際領域融合棟2階（当時）の皆様，ATR 脳情報解析研究所 動的脳イメージング研究室（当時）の皆様にはお世話になりました。また同研究所 計算脳イメージング研究室の皆様にも，本研究の初期段階においてアドバイスを頂くなど，ご協力をいただきました。博士号取得まで大変な遠回りをしてきましたが，当時の経験なしに本研究は形にできなかったと思っています。

現所属である本学バイオサイエンス研究科 駒井 章治先生，慶應義塾大学文学部 皆川 泰代先生には，博士号取得にご理解いただき暖かいサポートをいただきました。また，教務担当の情報科学研究科の安本 慶一先生や事務の皆様には，取得に当たっての相談などで大変お世話になりました。

学位取得はあくまで研究者としての第一歩だと思っています。ご協力いただいた皆様，見守ってくださった皆様の期待を裏切らぬよう，研究者として挑戦する哲学を忘れずに次の一歩へ臨んでまいります。

参考文献

- [1] Carol L Krumhansl. *Cognitive Foundations of Musical Pitch*. Oxford University Press, New York, 1990.
- [2] Stefan Koelsch, Thomas Gunter, Angela D. Friederici, and Erich Schröger. Brain indices of music processing: "nonmusicians" are musical. *Journal of Cognitive Neuroscience*, 12(3):520–41, 2000.
- [3] Martin A Rohrmeier and Stefan Koelsch. Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83(2):164–75, 2012.
- [4] Stefan Koelsch. Brain correlates of music-evoked emotions. *Nature reviews. Neuroscience*, 15(3):170–80, 2014.
- [5] Leonard B Meyer. *Emotion and Meaning in Music*. University of Chicago Press, Chicago, 1956.
- [6] David Temperley. *Music and Probability*. MIT Press, Cambridge, Massachusetts, 2007.
- [7] Walter Piston. *Harmony*. W.W.Norton & Company, New York, 1948.
- [8] Carol L Krumhansl and Roger N Shepard. Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 5(4):579–94, 1979.
- [9] Carol L Krumhansl, Jamshed J Bharucha, and Edward J Kessler. Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology. Human Perception and Performance*, 8(1):24–36, 1982.
- [10] Carol L Krumhansl and Edward J Kessler. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89(4):334–68, 1982.

- [11] Jamshed J Bharucha and Keiko Stoeckig. Reaction time and musical expectancy: priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, 12(4):403–10, 1986.
- [12] Jamshed J Bharucha and Keiko Stoeckig. Priming of chords: spreading activation or overlapping frequency spectra? *Perception & Psychophysics*, 41(6):519–24, 1987.
- [13] Emmanuel Bigand and Marion Pineau. Global context effects on musical expectancy. *Perception & Psychophysics*, 59(7):1098–107, 1997.
- [14] Jamshed J Bharucha. Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, 5:1–30, 1987.
- [15] Allan M Collins and Elizabeth F Loftus. A spreading-activation theory of semantic processing. *Psychological Review*, 82(6):407–28, 1975.
- [16] Barbara Tillmann, Jamshed J Bharucha, and Emmanuel Bigand. Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107(4):885–913, 2000.
- [17] Emmanuel Bigand, Bénédicte Poulin, Barbara Tillmann, François Madurell, and Daniel a D’Adamo. Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, 29(1):159–71, 2003.
- [18] Barbara Tillmann and Emmanuel Bigand. Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology. Human Perception and Performance*, 27(5):1185–96, 2001.
- [19] Barbara Tillmann and Géraldine Lebrun-Guillaud. Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research*, 70(5):345–58, 2006.

- [20] Barbara Tillmann and Frederic Marmel. Musical expectations within chord sequences: Facilitation due to tonal stability without closure effects. *Psychomusicology: Music, Mind, and Brain*, 23(1):1–5, 2013.
- [21] Risto Näätänen, Petri, Paavilainen, Teemu Rinne, and Kimmo Alho. The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12):2544–90, 2007.
- [22] Stefan Koelsch, Thomas C Gunter, Erich Schröger, Mari Tervaniemi, Daniela Sammler, and Angela D Friederici. Differentiating ERAN and MMN: an ERP study. *Neuroreport*, 12(7):1385–9, 2001.
- [23] Sakari Leino, Elvira Brattico, Mari Tervaniemi, and Peter Vuust. Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Research*, 1142:169–77, 2007.
- [24] Stefan Koelsch, Erich Schroger, and Thomas C Gunter. Music matters: preattentive musicality of the human brain. *Psychophysiology*, 39(1):38–48, 2002.
- [25] Burkhard Maess, Stefan Koelsch, Thomas C Gunter, and Angela D Friederici. Musical syntax is processed in Broca’s area : an MEG study. *Nature Neuroscience*, 4(5):540–5, 2001.
- [26] Stefan Koelsch. Significance of Broca’s area and ventral premotor cortex for music-syntactic processing. *Cortex*, 42(4):518–20, 2006.
- [27] Barbara Tillmann, Petr Janata, and Jamshed J Bharucha. Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, 16(2):145–61, 2003.
- [28] Carol A Seger, Brian J Spiering, Anastasia G Sares, Sarah I Quraini, Catherine Alpeter, James David, and Michael H Thaut. Corticostriatal contribu-

- tions to musical expectancy perception. *Journal of Cognitive Neuroscience*, 25(7):1062–77, 2013.
- [29] David Caplan, Nathaniel Alpert, Gloria Waters, and Anthony Olivieri. Activation of Broca’s area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*, 9(2):65–71, 2000.
- [30] Angela D Friederici, Shirley–Ann Rüschemeyer, Anja Hahne, and Christian J Fiebach. The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. *Cereb Cortex*, 13(2):170–7, 2003.
- [31] Aniruddh D Patel. Language, music, syntax and the brain. *Nature Neuroscience*, 6(7):674–81, 2003.
- [32] Luciano Fadiga, Laila Craighero, and Alessandro D’Ausilio. Broca’s area in language, action, and music. *Annals of the New York Academy of Sciences*, 1169:448–58, 2009.
- [33] Barbara Tillmann. Music and language perception: expectations, structural integration, and cognitive sequencing. *Topics in Cognitive Science*, 4:568–84, 2012.
- [34] David Kraehenbuehl and Edgar Coons. Information as a measure of the experience of music. *Journal of Aesthetics and Art Criticism*, 17(4):510–22, 1959.
- [35] Samer Abdallah and Mark Plumbley. Information dynamics: patterns of expectation and surprise in the perception of music. *Connection Science*, 21(2-3):89–117, 2009.
- [36] David Temperley. A probabilistic model of melody perception. *Cognitive Science*, 32(2):418–44, 2008.

- [37] Alexander Sheh and Daniel P W Ellis. Chord segmentation and recognition using EM-trained hidden Markov models. In *Proceedings of International Conference on Music Information Retrieval*, 183–9, 2003.
- [38] Hélène Papadopoulos and Geoffroy Peeters. Joint Estimation of Chords and Downbeats From an Audio Signal. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(1):138–52, 2011.
- [39] Dan Ponsford, Geraint Wiggins, and Chris Mellish. Statistical Learning of Harmonic Movement. *Journal of New Music Research*, 28(2):150–77, 1999.
- [40] Fred Lerdahl and Ray Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, Massachusetts, 1983.
- [41] WB De Haas, Martin Rohrmeier, RC Veltkamp, and Frans Wiering. Modeling harmonic similarity using a generative grammar of tonal harmony. In *10th International Society for Music Information Retrieval Conference*, 549–54, 2009.
- [42] Martin Rohrmeier. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53, 2011.
- [43] Marcus T Pearce and Geraint A Wiggins. Auditory expectation: the information dynamics of music perception and cognition. *Topics in Cognitive Science*, 4(4):625–52, 2012.
- [44] Daniela Sammler, Stefan Koelsch, and Angela D Friederici. Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*, 47(6):659–73, 2011.
- [45] David Butler. Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, 6(3):219–42, 1989.

- [46] Emmanuel Bigand, François Madurell, Barbara Tillmann, and Marion Pineau. Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1):184–97, 1999.
- [47] David Marr. *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press, Cambridge, Massachusetts, 1982.
- [48] P Read Montague, Peter Dayan, and Terrence J Sejnowski. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16:1936–47, 1996.
- [49] Wolfram Schultz, Peter Dayan, and P Read Montague. A neural substrate of prediction and reward. *Science*, 275:1593–9, 1997.
- [50] John P O’Doherty, Peter Dayan, Karl Friston, Hugo Critchley, and Raymond J Dolan. Temporal Difference Models and Reward-Related Learning in the Human Brain. *Neuron*, 38(2):329–37, 2003.
- [51] John P O’Doherty, Alan Hampton, and Hackjin Kim. Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104(1):35–53, 2007.
- [52] Dominic J Barraclough, Michelle L Conroy, and Daeyeol Lee. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, 7(4):404–10, 2004.
- [53] Wako Yoshida and Shin Ishii. Resolution of uncertainty in prefrontal cortex. *Neuron*, 50(5):781–9, 2006.
- [54] David Badre and Bradley B Doll, Nicole M Long, and Michael J Frank. Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, 73(3):595–607, 2012.

- [55] Maël Donoso, Anne G E Collins, and Etienne Koechlin. Foundations of human reasoning in the prefrontal cortex. *Science*, 344(6191):1481–6, 2014.
- [56] Jan Gläscher, Nathaniel Daw, Peter Dayan, and John P O’Doherty. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron*, 66(4):585–95, 2010.
- [57] Nathaniel D Daw and Kenji Doya. The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, 16:199–204, 2006.
- [58] Martijn J Mulder, Leendert van Maanen, and Birte U Forstmann. Perceptual decision neurosciences - A model-based review. *Neuroscience*, 277:872–84, 2014.
- [59] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, Massachusetts, 1998.
- [60] Roger N Shepard. Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, 36(12):2346–53, 1964.
- [61] Joseph B Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.
- [62] Stephen C Johnson. Hierarchical clustering schemes. *Psychometrika*, 32:241–54, 1967.
- [63] B Tillmann, Emmanuel Bigand, and Marion Pineau. Effects of global and local contexts on harmonic expectancy. *Music Perception*, 16(1):99–117, 1998.
- [64] Hirotugu Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–23, 1974.

- [65] Katy Noland and Mark B Sandler. Key Estimation Using a Hidden Markov Model. In *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, 121–6, Victoria, BC, Canada, 2006.
- [66] Matthias Mauch and Simon Dixon. Simultaneous estimation of chords and musical context from audio. *IEEE Transactions on Audio, Speech and Language Processing*, 18(6):1280–9, 2010.
- [67] David Huron. *Sweet Anticipation: Music and the Psychology of Expectation*. A Bradford Book. MIT Press, Cambridge, Massachusetts, 2006.
- [68] Edward J Kessler, Christa Hansen, and Roger N Shepard. Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, 2:131–65, 1984.
- [69] Carol L Krumhansl, Gregory J Sandell, and Desmond C Sergeant. The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, 5(1):31–77, 1987.
- [70] Emmanuel Bigand and Bénédicte Poulin-Charronnat. Are we ”experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100(1):100–30, 2006.
- [71] Konrad P Körding and Daniel M Wolpert. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10(7):319–26, 2006.
- [72] Wendy J Adams, Erich W Graf, and Marc O Ernst. Experience can change the ’light-from-above’ prior. *Nature Neuroscience*, 7(10):1057–8, 2004.
- [73] Makoto Miyazaki, Shinya Yamamoto, Sunao Uchida, and Shigeru Kitazawa. Bayesian calibration of simultaneity in tactile temporal order judgment. *Nature Neuroscience*, 9(7):875–7, 2006.

- [74] Ken-ichi Sawai, Yoshiyuki Sato, and Kazuyuki Aihara. Auditory time-interval perception as causal inference on sound sources. *Frontiers in Psychology*, 3:524, 2012.
- [75] Michael D Lee. A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, 30(3):1–26, 2006.
- [76] Etienne Koechlin. An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369(1655):20130474, 2014.
- [77] Makoto Miyazaki, Hiroshi Kadota, Kozue S Matsuzaki, Shigeki Takeuchi, Hirofumi Sekiguchi, Takuo Aoyama, and Takanori Kochiyama. Dissociating the neural correlates of tactile temporal order and simultaneity judgements. *Scientific Reports*, 6:23323, 2016.
- [78] Izzet B Yildiz, Katharina von Kriegstein, and Stefan J Kiebel. From Bird-song to Human Speech Recognition: Bayesian Inference on a Hierarchy of Nonlinear Dynamical Systems. *PLoS Computational Biology*, 9(9), 2013.
- [79] Martin J Pickering and Simon Garrod. An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences*, 36(4):329–47, 2013.
- [80] William F Thompson and Lola L Cuddy. Sensitivity to key change in chorale sequences: A comparison of single voices and four-voice harmony. *Music Perception*, 7:151–68, 1989.
- [81] Nikolaus Steinbeis, Stefan Koelsch, and John a Sloboda. The role of harmonic expectancy violations in musical emotions: evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18:1380–93, 2006.

- [82] Stefan Koelsch, Thomas Fritz, and Gottfried Schlaug. Amygdala activity can be modulated by unexpected chord functions during music listening. *Neuroreport*, 19(18):1815–9, 2008.

付録

A. 業績リスト

査読付学術論文

1. Satoshi Morimoto, Gerard B. Remijn, Yoshitaka Nakajima, Computational-model-based analysis of context effects on harmonic expectancy, *PLOS ONE*, 11, 3, e0151374, 2016.3, 第3章及び第4章に対応

査読付国際会議発表

1. Satoshi Morimoto, Gerard B. Remijn, Yoshitaka Nakajima, Computational model-based analysis of context effects on chord processing, *13th International Conference on Music Perception and Cognition*, 2014.8, 第3章に対応

B. 補足データ

B.1 参加者の個人差について

実験参加者の内的な期待の個人差について確認するため、いくつかの側面からDORと提案モデルの推定された期待行列について調べた。まず図27(a)に実験で用いた連鎖和音刺激に対する平均DORの値と、対応する標準偏差の関係を散布図で表す。高いDORを回答した連鎖和音に対するばらつきが小さい傾向が見られた。同様にBUモデルの期待行列に含まれる和音進行について、平均値と標準偏差の関係を図27(b)に示す。Iの和音からの和音進行を除けば、高い期待を示す和音進行に対してばらつきが小さくなる傾向が得られた。DORや期待行列は中間値である5、或いは0.5を境として適切と不適切が分かれている。そのため、内的により判断が明快となる両端の値に近づくにつれ、標準偏差が小さくなる可能性がある。しかし、似た平均値を示す条件において標準偏差の値自体もばらついており、個人差が存在する可能性も捨てきれないと言える。

さらに音楽の経験値が期待のパターンに及ぼす影響について調べるため，個人ごとに推定した BU モデルの期待行列について，個人間のユークリッド距離行列を求めて最長距離法による階層クラスタリングを行った（図 28）．今回の実験参

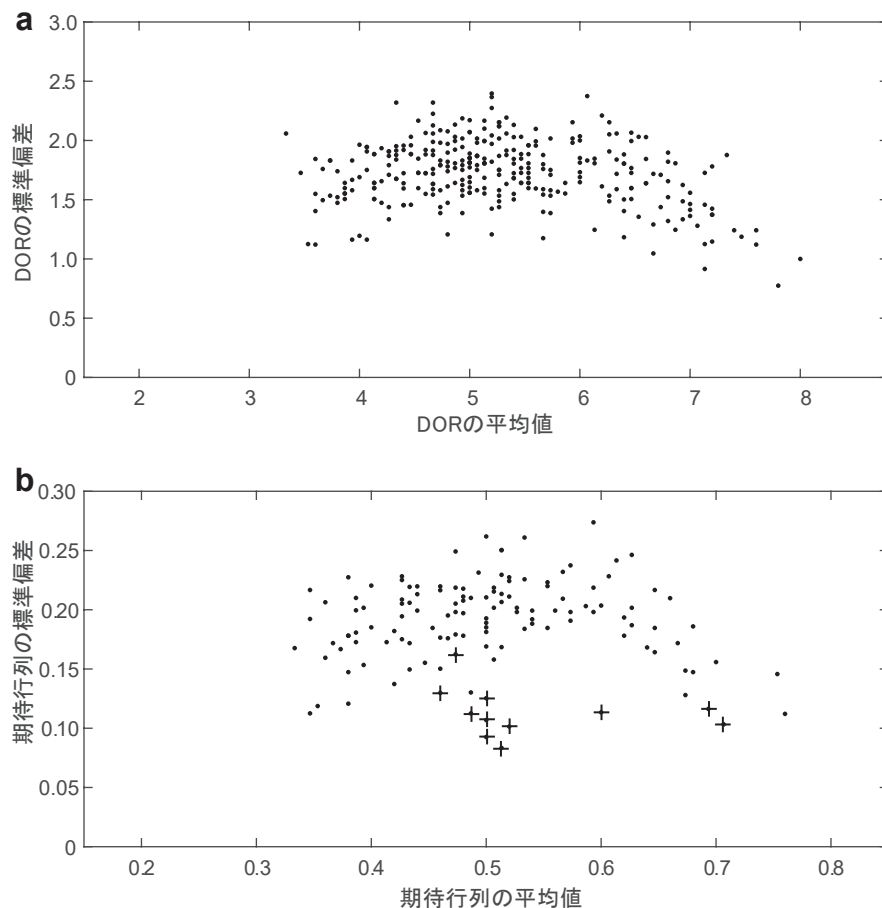


図 27 DOR 及び期待行列における実験参加者間のばらつき

(a) 全参加者の平均 DOR に対する標準偏差と，(b)BU モデルの推定期待行列の平均値に対する標準偏差の関係を散布図で表す．図中の各点は連鎖和音または和音進行を表す．ただし，(a) は同じ和音への遷移を含む連鎖和音を除外してプロットした．また (b) における + 印は，期待行列において I の和音からの和音進行を表す（図 26(b) 及び本文中の議論参照）．

加者内では、音楽経験による明確な影響は見受けられなかった。先行研究においても、音楽経験の長さが及ぼす影響は比較的小さいことが報告されており [70], 本結果と矛盾ない。しかし、評価尺度の取り方に個人差があると考えられるため、期待の大きさについてのみ評価できていない可能性が高い。また、音楽経験年数は音楽が至る場面においても氾濫している現代社会において、必ずしも楽曲への曝露の目安とはならない問題がある。

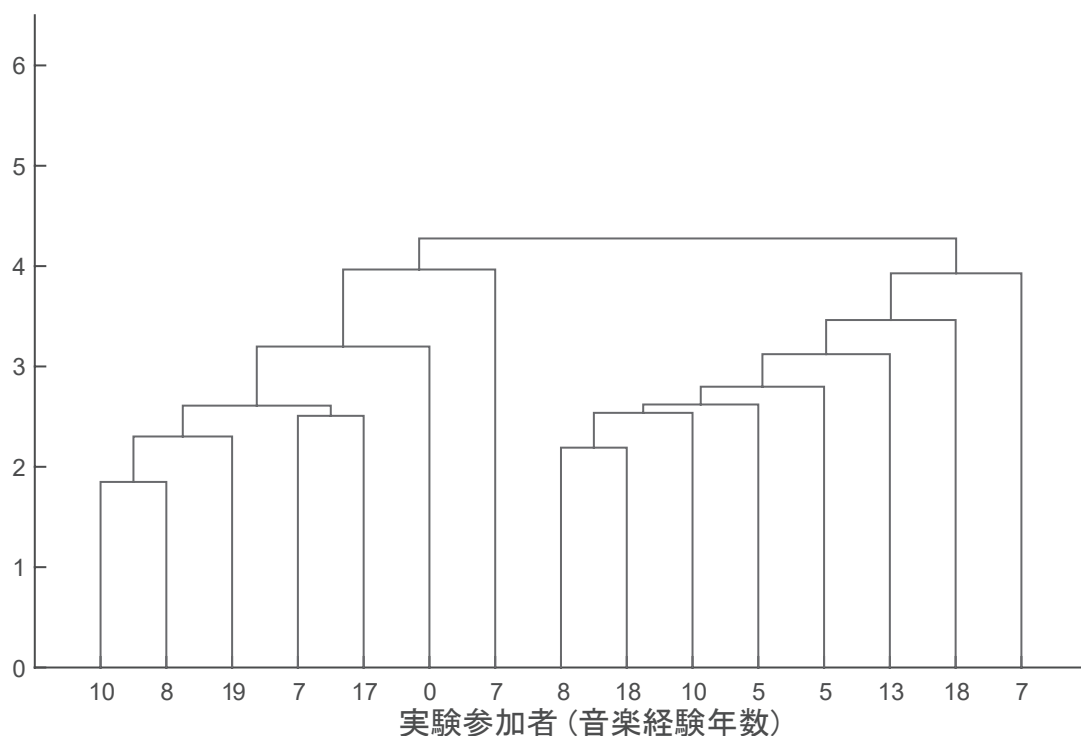


図 28 BU モデルの期待行列への音楽経験の影響

推定された期待行列の階層クラスタリングによる樹形図を示す。実験参加者を標識している数字は、音楽経験年数を表す。ただし、内的な期待の大きさだけでなく評価尺度の取り方の違いも反映されていることに注意が必要である。

B.2 参加者個人内のDORのばらつきについて

個人内のDORのばらつきについて見積もるため、本実験において最も単純な刺激である2連鎖条件における音程の効果について調べた。図29に全実験参加者の各音程に対するDORの平均値と標準偏差を示す。2連鎖条件のような比較的単純な課題においても、ばらつきがあることが分かる。また同じ和音が連続する刺激（図中、音程が0のとき）に対するDORは、参加者によって明らかに回答の傾向が異なっている。実験後の内観調査においても、回答に対する混乱を報告する参加者が多かった。

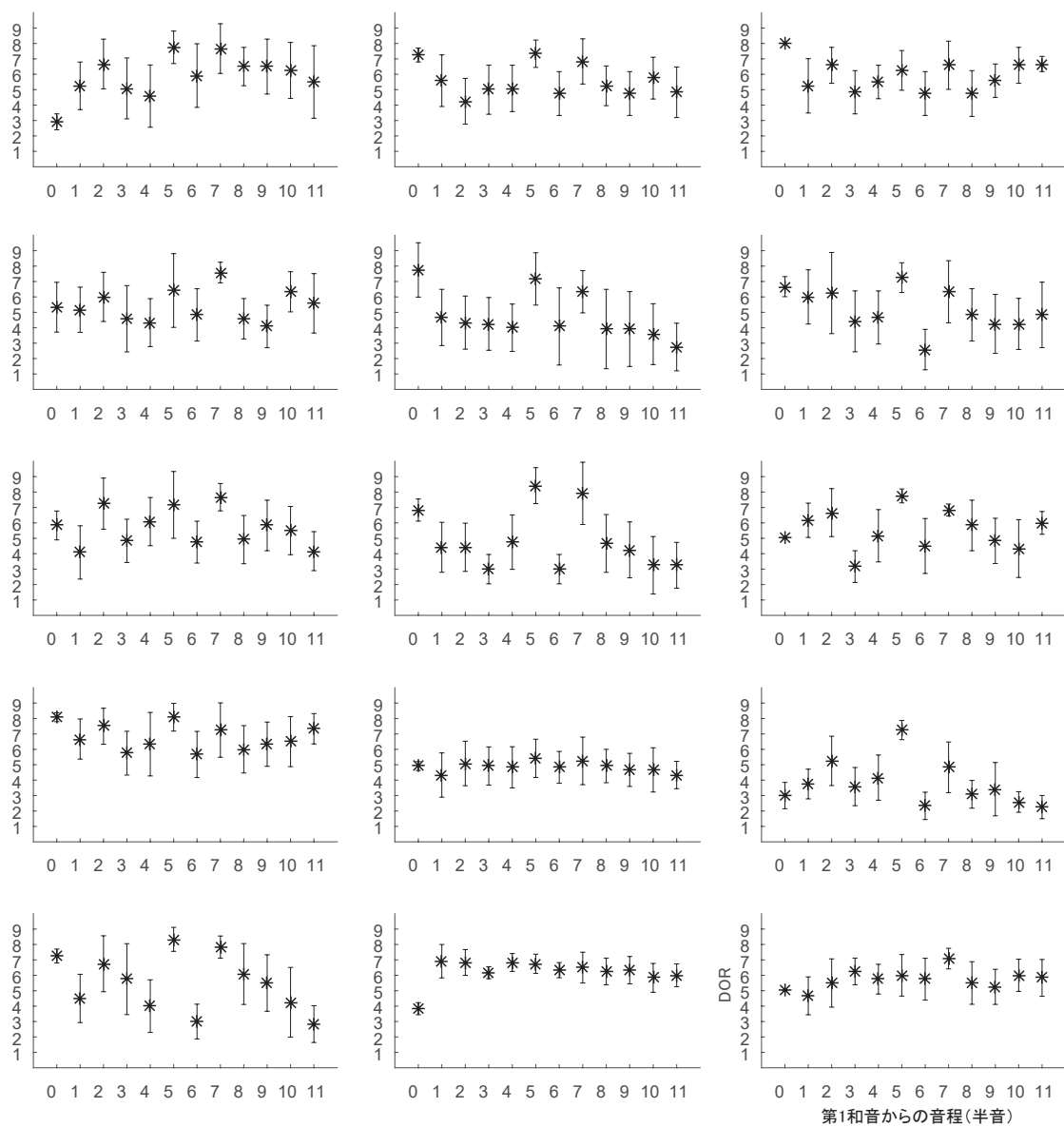


図 29 2連鎖条件における実験参加者ごとのDORのばらつき
 15名の参加者の2連鎖条件におけるDORについて、第1和音から第2和音への音程に着目したときの平均DOR(*印)と標準偏差を表す。

B.3 音楽的期待と音程の基準の関係

音楽的期待と調の関係について調べるため、内的な音程の基準（調）の分布のエントロピーを算出し、推定されたBUモデルにおける期待（適切度）の大きさとの関係をプロットした（図30）。2連鎖条件，3連鎖条件，4連鎖条件それぞれ相関係数は $r = -0.11$ ， $r = -0.54$ ， $r = -0.42$ であり，期待と音程の基準のエントロピーの間に一定の相関がみられた。

音楽理論上，調の確立はIの和音に対する終止感の強さに関与すると考えられ，今後モデルを検証する心理実験として検討していく必要がある。

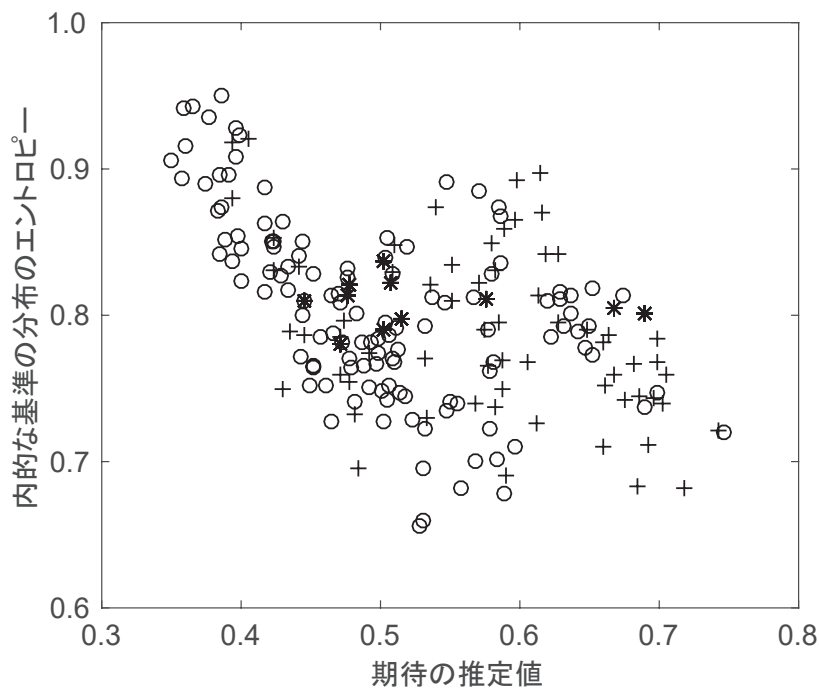


図 30 BUモデルにおける期待と音程の基準の関係

推定されたBUモデルにおける，期待（横軸）と音程の基準 X のエントロピーの関係を散布図で表す．* は2連鎖条件， \circ は3連鎖条件，+ は4連鎖条件を示す。