

音高ずれ検知の旋律依存性

白土 保*, 柳田 益造**

郵政省通信総合研究所*, 同志社大学**

*shirado@crl.go.jp**, *myanagid@doshisha.ac.jp***

聴取者が単旋律の音高を評価する際のモデルの基本的枠組みを提案し、モデルの一部「音高ずれ検知の旋律依存性」仮説を検証している。この仮説は、音高ずれの検知されにくさの指標を確率的な特徴量で定義するとともに、旋律の複雑さを跳躍進行と非全音階音の出現頻度によって定義した上で、「音高ずれの検知されにくさの度合は、その音を含む旋律の複雑さの度合に比例する」というものである。音大生を被験者とし複雑さが異なる3種類の旋律を用いた聴覚実験を行った結果、完全とは言えないまでも仮説の妥当性がほぼ確認されている。

Dependency of off-scale sensation on complexity of melodies

Tamotsu Shirado*, Masuzo Yanagida**

Communications Research Laboratory, Doshisha University***

Discussed in the present paper is off-scale sensation on tones in melodies. Experimental result shows that listeners can perceive off-scale sensation on a tone in simple melodies easier than that in complex melodies.

1 まえがき

我々は、演奏の評価に関する工学モデルの構築に取り組んでいる。演奏の評価には様々なものがあるが、一般に「音楽性の評価」は聴取者の感性や主観に大きく依存するため、工学的に取り扱いにくい。そこで現在のところ、比較的普遍性のあると考えられる「音高の評価」に限って工学モデルの構築を行っている。

単音の音高評価については、音の周波数弁別閾に関する研究が過去多くなされてきており、その周波数特性などが詳細に調べられてきている¹⁾。また、音程の弁別に関する研究については、音程の範疇知覚などの事実が明らかにされている²⁾。しかし、単旋律の音高評価に関して、周波数弁別の観点から調べた研究はまだない。単旋律の音高評価においては、単音に対する周波数弁別閾や音程についての範疇検知の他、音楽的文脈などのさまざまな要因が関係していると思われる。

今回は音高の評価モデルの基本的枠組みを提案し、モデルの一部として「単旋律における音高ずれの検知されにくさの度合は、その音を含む旋律の複雑さの度合に比例する」、という仮説を立て、音大生を被験者とした聴覚実験を行うことにより仮説の妥当性を検討した。

2 音高ずれの検知

フレットのない弦楽器などによる単旋律演奏の音高評価においては、たとえ絶対音感を持つ聴取者であっても演奏音の絶対的高さを一音ずつ評価しているとは考えられない。なぜなら、演奏者が調律の際に基準とする音の音高（例えばA₄の周波数）が異なる状況（演奏者や曲）においても同じとなる保証は全くないが、この音高がどのような値であっても、絶対音感をもつ聴取者が演奏音を奇異に感じることはないからである。音高ずれの検知は、音の絶対的高さではなく、聴取者が持つ音律感覚と、演奏音聴取を通じて聴取者が確立した基準音の音高に基づい

ている、と考えられる。ただし、どの音名の音を基準音にするかは被験者や聴取の状況により異なると考えられる。聴取者はこの音律及び基準音の音高に基づき、音名をつけることのできる音として持つべき基本周波数（ここでは期待周波数と呼ぶ）をすべての音名の音に割り当てる。そして、聴取した音の基本周波数と、それに最も近い期待周波数の比較により音高評価が行われる、と考えられる。以上の考察に基づき、単旋律における音高評価の工学的モデルの基本的枠組みを提案する。

【単旋律音高評価モデル】

演奏者 P (Playerの頭文字) が音符の並び $\Phi_i = \phi_1, \phi_2, \dots, \phi_i$ からなる譜面を演奏する際、時刻 t の演奏音を $\omega_P(t)$ (演奏開始時刻を $t = 0$)、音符 ϕ_i に対する演奏開始及び終了時刻を b_i 及び e_i としたとき、 ϕ_i に対する演奏音 ω_{P_i} を次式で定義する。

$$\omega_{P_i} \triangleq \omega_P(t), \quad b_i < t < e_i \quad (1)$$

演奏音は、任意の時刻において単音を高々一つ含むものと仮定する。

以下 ω_{P_i} を単に ω_i と記す。各 ω_i は既定常と仮定する。音列 $\omega_1, \omega_2, \dots, \omega_i$ を Ω_i と記す。ここで ω_i の基本周波数を $f_0(\omega_i)$ 、聴取者を L (Listenerの頭文字) としたとき、 $f_l(L, i) \leq f_0(\omega_i) \leq f_h(L, i)$ が満たされるような $f_0(\omega_i)$ に対しては「音高が正確」と感じ、満たされない場合は「音高がずれている」と感じるものとする。実際にはずれの検知は確率的であるため、 $f_l(L, i)$ 及び $f_h(L, i)$ は、ある一定のずれ検知確率に対して決められる(値については後述する)。 $f_l(L, i)$ 及び $f_h(L, i)$ を、それぞれ式(2)及び式(3)で定義する。

$$f_l(L, i) \triangleq f_E(T(L), B(L, \Omega_{i-1}), n(\phi_i)) \\ \times l(L, s(k(L, \Omega_{i-1}), n(\phi_i)), \Phi_i) \quad (2)$$

$$f_h(L, i) \triangleq f_E(T(L), B(L, \Omega_{i-1}), n(\phi_i)) \\ \times h(L, s(k(L, \Omega_{i-1}), n(\phi_i)), \Phi_i) \quad (3)$$

ここで、

$f_E(T, B, n)$: 音律 T , T の基準の周波数 B における絶対音名 n の音の基本周波数 (期待周波数) .

l : f_E に対する f_l の周波数比.

h : f_E に対する f_h の周波数比.

$T(L)$: 聴取者 L が許容する音律.

$B(L, \Omega)$: 聴取者 L が音列 Ω を聞いた時点における音律 T の基準の周波数.

$n(\phi)$: 音符 ϕ の絶対音名.

$s(k, n)$: 調 k における音名 n の階名.

$k(L, \Omega)$: 聴取者 L が音列 Ω を聞いた時点で感じる調.

周波数比 l 及び h は以下に述べる「ずれ弁別閾」と「ずれ感度」の二つの特徴量の関数と仮定する.

[ずれ弁別閾及びずれ感度]

周波数比 $f_0(w_i)/f_E$ をセント値で表したものを「ずれ値」と呼ぶことにする. ここで、正のずれ値を持つ音に対し「高くずれている」と感じる割合及び負のずれ値を持つ音に対し「低くずれている」と感じる割合を「正反応率」(逆の場合における反応率を「負反応率」と呼ぶ. このとき正反応率 P は、ogive (累積正規分布) に近い関数で求められると期待される³⁾).

図 1 に示すように ogive 曲線は、特徴量 D (縦軸上の値 {正反応率} が 0.5 となる点の、横軸上の値 {刺激量}) 及び S (その点における曲線の傾き) で代表することができる. 本論文では、 D を「ずれ弁別閾」、 S を「ずれ感度」と呼ぶ.

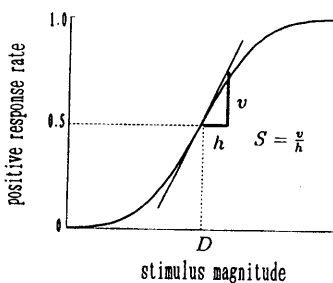


図 1 ずれ弁別閾及びずれ感度

正反応率 $P(0.5 < P \leq 1.0)$ に対し、周波数比 $l = f_l/f_E$ 及び $h = f_h/f_E$ を決める. P をこの範囲に限定した理由は、チャンスレベル (0.5 以下) 以下の確率は、ずれの検知にとってあまり意味がないと考えられるからである. P の具体的な値はケースバイケースで決めることができるが、ここでは特定しない.

このとき $1/l$ 及び h は D が大きく S が小さいほど大きい値とるが、このことを「音高ずれが検知されにくい」と表現する.

関数 l 及び h の 3 番目の引数 Φ_i は音楽的文脈を表している. ここでは、音楽的文脈を旋律の複雑さの観点でとらえる. ただし「旋律の複雑さ」の度合は、(1) 旋律中で隣合った音のすべての音程のうち、跳躍進行 (音楽理論では、短 3 度以上の進行と定義される) の相対出現頻度 J (Jump の頭文字)、及び (2) 旋律中の非全音階音の相対出現頻度 C (Chromatic の頭文字) に比例する関数の値として決められるものとする.

ここで以下の仮説を立てる.

仮説 音高ずれの検知されにくさの度合は、その音を含む旋律の複雑さの度合に比例する.

この仮説は、以下 3 つの条件のいずれかが満たされれば成立する. 旋律の複雑さの度合に対し、(条件 i) D は比例し S は反比例する、(条件 ii) D は比例し S は依存しない、(条件 iii) S は反比例し D は依存しない.

仮説を立証するためには、旋律の複雑さが異なる複数種類の旋律間での D 及び S の差を調べ、それが先に述べた 3 つの条件のどれかに落ちることを示せば良い. 今回は、3 種類の旋律を用いて仮説検証のための実験を行った.

今回の実験においては、試行中常に f_E がほぼ既知とみなせる条件のもとに D 及び S を測定している. つまり今回は式 (2), (3) の関数 l 及び h の性質についてのみ調べていることになる.

3 音高ずれの検知に関する聴覚実験

[被験者]

平均律を許容する9名の音大生(付録参照)。

[実験刺激に用いた旋律]

実験刺激は、図2に示す旋律1(全音階音列)、旋律2(半音階音列)及び旋律3(12音音楽のセリーの一部)を原旋律として用い、派生させた音列である。これらの旋律の複雑さを前述の定義により評価すると、旋律1は $J=0, C=0$ 、旋律2は $J=0, C \approx 0.4$ 、旋律3は $J \approx 0.9, C \approx 0.4$ となるため、旋律1<旋律2<旋律3の順に複雑さが増すことになる。



図2 実験刺激に用いた旋律

[刺激音の詳細と呈示スキーマ]

音列を構成する各音の周波数は、 $A_4=440+\alpha\text{Hz}$ を基準として与えた。ここで α は、 $-18.5 \sim +19.5$ (即ち、 $A_4=440\text{Hz}$ に対する周波数比が -75 セント $\sim +75$ セント)の範囲からランダムに選んだ値とした。ゲタ α をはかせた理由は、直前の試行で聴取した音列の音高感が現在の試行に影響しないようにするためである。ただし、↓のついた対象音(f a, s o l)のうちいずれかひとつの周波数は、この周波数から更に β セントずらせた値とした(周波数をずらせた対象音

を「ずれ音」と呼ぶ)。 β のバリエーションは、 $\beta = 0, \pm 5, \pm 10, \dots, \pm 50$ (21通り)とした。つまり旋律、対象音及びずれ値 β についてのバリエーションとして $3 \times 2 \times 21 =$ 計126種類の音列を用いた。実験には同一の β を持つ音列を6回ずつ用いた。従って実験に用いた音列の総数は、 $126 \times 6 = 756$ 個である。

このようにして生成された756個の刺激音列は、ランダムな順番で一つずつ呈示された。1試行の時間的スキーマを図3に示す。各試行では、刺激音列の前にダミー音(刺激音列のcue)を挿入した。

実験音としては、音の持続時間300msec、立ち上がり時間30msec、立ち下がり時間50msecの調波合成音を用いた。実験音は、あらかじめDAT(Digital Audio Tape)に録音され、通常の教室でヘッドホン(STAX A-PRO)を通して被験者に呈示された。

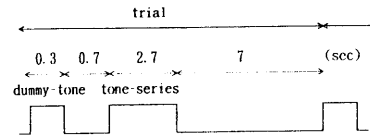


図3 呈示スキーマ

[回答方法]

被験者は、音列中の特定の音に注目して音列を聴取し、注目している音の音高に関して以下の選択肢のいずれか一つを選ぶよう求められた。選択肢は、「高くずれているように感じる」/「低くずれているように感じる」/「合っているように感じる」/「ずれているように感じるが高低は不明」/「ずれているか否か判断不可能」とした。音列中の何番目の音を判定の対象とするかの指示は、回答用紙に試行毎に記しておいた。判定の対象とする音としては、ずれ音に注目するよう指示した。また被験者には、実験に用いた旋律のすべての譜面をいつでも参照できるようにした。

実験は34ラウンドに分けて2日間で行った。1ラウンド5分、ラウンド間には15分程度の休み時間をおいた。

本実験で用いたすべての音の基本周波数は、490Hz～1.1kHzの範囲にある。この周波数範囲にある基本周波数を持つすべての音の、単音としての周波数弁別閾はほぼ均一である¹⁾。従って本実験では、刺激音列相互の絶対音高の違いによる影響は小さいと考えられる。

本論文では、各音列の開始の音をd oとして、階名で各音を指すことにする。これは、刺激音列における各音の高さがA₄=440Hz平均律で与えられる周波数から一様に最大±75セントだけシフトしており、音を指すのに絶対音名を用いると、半音違いの音との区別の問題が生じるからである。

4 実験結果

前述したように実験では、旋律、ずれ音およびずれ値のバリエーションについて同一の刺激音列が6回ずつ用いられている。各条件における6回の試行中「低くずれている」という回答の数をc_l、「高くずれている」という回答の数をc_h、「ずれているか否か判断不可能」以外の回答の数をc_oとし、各旋律、及び各ずれ音に対してP_l⁺=c_l/c_o及びP_h⁺=c_h/c_oを評価した。lはβ<0では正反応の数、β>0では負反応の数を表す。hはβ<0では負反応の数、β>0では正反応の数を表すのでP_lはβ<0で正反応率(P_l⁺と記す)、β>0で負反応率(P_l⁻と記す)を表すことになる。またP_h⁺,P_h⁻は、この逆の場合の量を表す。

4.1 反応率の被験者間平均及び標準偏差

各条件における被験者9名のP_l(平均を▽、標準偏差を実線で表示)及びP_h(平均を△、標準偏差を点線で表示)をまとめてプロットしたものを、図4に示す。横軸はずれ値(単位はセント)、縦軸は反応率(正反応率あるいは負反応率)であ

る。図4は6つの図からなり、それぞれ旋律3種類×ずれ音2種類の計6つの条件に対応する。6つのいずれの図においても、ずれ値の絶対値に関して、P_l⁺及びP_h⁺の平均は単調増加の傾向が見られ、ほぼ直線で近似できる。この直線は、先に述べたogive曲線の第一次近似にあたる。

4.2 ずれ弁別閾及びずれ感度

P_l⁺の値は、呈示されたずれ音の基本周波数が期待周波数より低い場合における音高ずれの検知を、P_h⁺の値は、その逆の場合を表しているの以下、前者を低ずれ反応、後者を高ずれ反応と呼ぶことにする。

6つの条件それぞれにおいて、0～-50セント(11通り)のずれ値に対する被験者全員の低ずれ反応P_l⁺のデータを用い、ずれ値をΔとして直線P=a+bΔへの回帰分析を行った。即ち回帰分析に用いるデータ対(x_i,y_i)としては、(Δ_i,P_l⁺),i=1..Nを用いた。ここでN=99(ずれ値のバリエーション11通り×被験者9人)である。b及びaはそれぞれ式(1),(2)で推定し、ずれ感度S=b、ずれ弁別閾D=(0.5-a)/bを求めた。

$$b = \frac{\sum_{i=1}^N x_i y_i - N \bar{x} \bar{y}}{\sum_{i=1}^N x_i^2 - N \bar{x}^2} \quad (4)$$

$$a = \bar{y} - b \bar{x} \quad (5)$$

ここで、 $\bar{x} = \sum_{i=1}^N x_i / N$, $\bar{y} = \sum_{i=1}^N y_i / N$.
ずれ感度Sの標準偏差σ_Sは次式、

$$\sigma_S = \sqrt{\frac{\sum_{i=1}^N (y_i - (a + b x_i))^2}{(N-2)(\sum_{i=1}^N x_i^2 - N \bar{x}^2)}} \quad (6)$$

また、ずれ弁別閾Dの標準偏差σ_Dは次式で推定した。

$$\sigma_D = \sqrt{\frac{\sum_{i=1}^N (x_i - \frac{y_i - a}{b})^2}{N-2}} \quad (7)$$

高ずれ反応P_h⁺についても同様に、0～+50セント(11通り)のずれ値に対する被験者全員のP_h⁺のデータを用いてD,σ_D,S,σ_Sを推定した。

推定結果を、図5(a),(b)に示す。両図とも▽、△はそれぞれ低ずれ反応、高ずれ反応における推定値を、縦棒は推定値の標準偏差を表す。この図を見ると、旋律1及び旋律2においては、階名やずれの方向によらず D 及び S の期待値はほぼ同じ値をとっていることがわかる。また旋律3における D の期待値は旋律1, 2に比べ、いずれの階名/ずれの方向においても一定値(20セント程度)以上大きい値をとっている。更に、旋律3における S の期待値は、旋律1, 2に比べいずれの階名やいずれのずれの方向においても一定値(0.007程度)より小さい値をとっていることがわかる。これらに有意な差があるか否かを以下の手続きで検定した。

4.3 ずれ弁別閾及びずれ感度に関する有意差検定

ずれ弁別閾の推定値は、平均値 D 、分散 σ_D^2 の正規分布に従うと仮定した。また一般に、サンプル数 N の直線回帰によって求められた直線の傾きの期待値を b 、その標準偏差を σ_b としたとき、傾きの推定値を表す確率変数 ψ についての式 $(\psi - b)/\sigma_b$ は、自由度 $N - 2$ の t 分布に従うが、 N が30以上の t 分布は、平均0、分散1の正規分布で置き換えても差し支えない。今回の実験では、 $N = 99$ であるのでこの条件を満たす。したがって、ずれ感度の推定値は平均値 S 、分散 σ_S^2 の正規分布に従う。以上をふまえ、旋律、ずれ音、ずれの方向のさまざまな条件間で D 及び S の有意差検定を行った。その結果を以下に示す。

[結果1] ずれ弁別閾及びずれ感度は、旋律1と旋律2の間では差がない(危険率1%)。

[結果2] ずれ弁別閾は、旋律1及び旋律2における値の方が旋律3における値より小さい。(危険率5%、弱い示唆)。

[結果3] ずれ感度は、旋律1及び旋律2における値の方が旋律3における値より大きい(危険率1%)。

[結果4] ずれ弁別閾及びずれ感度は、ずれの方向には依存しない(危険率5%)。

[結果5] ずれ弁別閾及びずれ感度は、階名 $f a$ と階名 $s o l$ の間では差がない(危険率1%)。

5 考察

[反応感度に関して]

図4におけるいくつかの図では、正反応率の最少値は必ずしも0に収束するようには見えない。この原因としては、被験者の反応が過敏になっていた可能性が指摘される。なぜなら被験者は、注目するように指示された音がずれているのではないかと、常に疑いながら回答したと思われるからである。したがって、本実験で得られたずれ弁別閾及びずれ感度の値は、一般的な聴取の状況における値より厳しい値となっている可能性がある。

[機能音声に関して]

ずれ弁別閾 D 及びずれ感度 S が、階名 $f a$ と階名 $s o l$ の間で差がなかったことは、必ずしも D 及び S が階名のもつ機能音声上の重要度に依存しないことを示唆するものではない。なぜなら $s o l$ は $d o$ から完全5度上の音であるが、一方 $f a$ はオクターブ上の $d o$ からは完全5度下の音程であり、 $f a$ と $s o l$ の機能音声上の重要度に大きな差があるとは必ずしも言えないからである。

[旋律の複雑さの定義の妥当性]

旋律1と旋律2の間でずれ弁別閾 D 及びずれ感度 S の差がなかった(結果1)理由としては、以下の3つの可能性がある。(1) D 及び S が旋律中の非全音階音の相対出現頻度 C には依存しない、(2) C よりも旋律中の跳躍進行の相対出現頻度 J への依存度が大きい、(3) J で進行の複雑さを表すのは不十分(旋律における音高の予測しにくさの観点からは、全音階的進行の方が半音階的進行より複雑である可能性がある)。

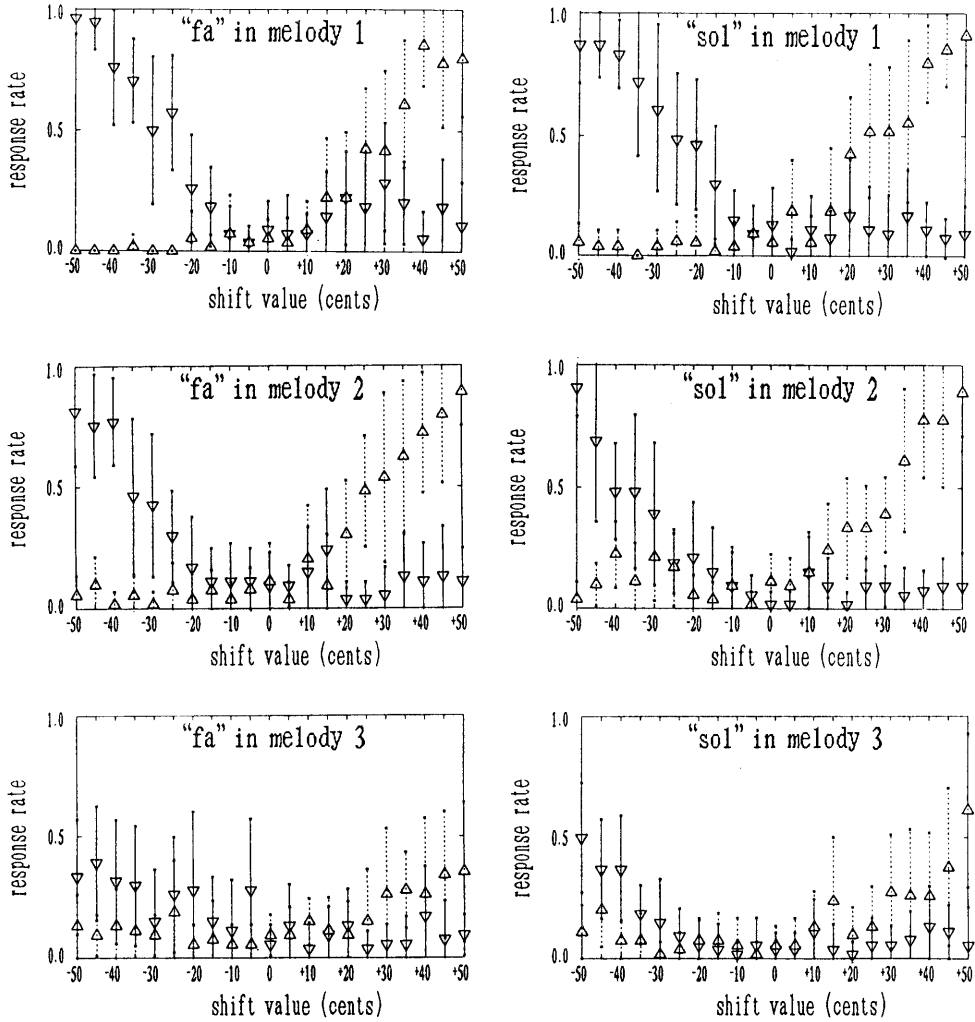
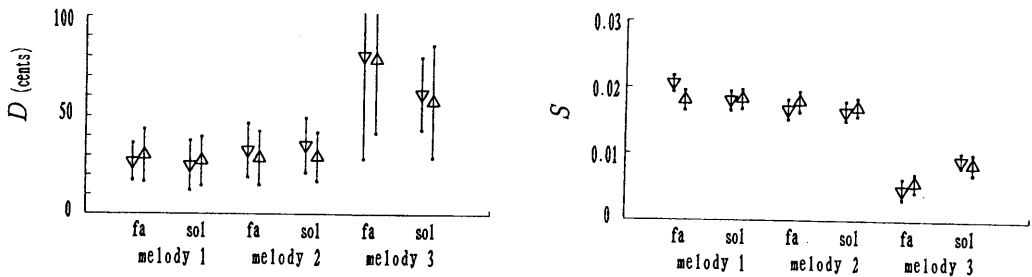


図4 反応率の被験者間平均及び標準偏差 (▽: P_l , △: P_h).



(a) ずれ弁別閾

(b) ずれ感度

図5 ずれ弁別閾及びずれ感度 (▽: 低ずれ反応, △: 高ずれ反応)

以上から、結果 1 は仮説を積極的に支持はしないが反証するものではない、また結果 2 及び結果 3 は仮説を支持する。ただし比例関係まで支持する結果は得られなかった。よって本実験により、完全とは言えないまでも仮説がほぼ立証された（これは仮説が成立するための条件 i に該当する）。

6 むすび

音高の評価モデルの基本的枠組みを提案し、モデルの一部として「単旋律における音高ずれの検知されにくさの度合は、その音を含む旋律の複雑さの度合に比例する」、という仮説を立てた。聴覚実験による検証を行った結果、仮説をある程度支持する結果が得られた。更に音高ずれの検知されにくさの度合は、聴取者の持つ音律からのずれの絶対値に依存し、ずれの方向には依存しないことが示唆された。

今後の課題としては、比例関係の詳細な検証、及び旋律の複雑さを与える関数の精密化が挙げられる。また音高評価モデルの内、機能と声上の重要度依存性の検証及び音律の基準 B の演奏音列 Ω 依存性については今回検討しなかった。今後は、これらを含め定量的な音高評価モデルの構築を目指す。

文献

- 1) B.C.J. Moore (大串健吾監訳), 聴覚心理学概論 (誠信書房, 東京, 1994), pp.172-178.
- 2) D.Deutsch 編 (寺西立年, 大串健吾, 宮崎謙一監訳), 音楽の心理学 (上) (西村書店, 新潟, 1987), pp.307-317.
- 3) 大山正, 今井省吾, 和気典二編, 感覚・知覚心理学ハンドブック (誠信書房, 東京, 1994), p.19.

予備実験 (平均律を許容する被験者の選別)

実験に参加した被験者は本来、声楽専攻 9 名及びピアノ専攻 3 名, 合計 12 名の音大生である。しかし今回の実験では、刺激音列作成の際に音律として平均律を用いているため、平均律を許容する被験者のみを用いることが必要となる。予備実験では、すべての音の基本周波数が平均律に基づいた音列を被験者に呈示し、各音がずれているように聞こえるか否かの反応 (false alarm 反応) を見ることによって平均律を許容する被験者を選別した。

実験刺激は、図 2 に示す旋律 1 及び旋律 2 (本実験と同じ旋律) を原旋律とし、音列を構成する各音の周波数を、 $A_4=440+\alpha\text{Hz}$ を基準に生成した音列の集合である。 α は本実験におけるものと同じ。刺激音列としては、旋律 1 からは 36 個、旋律 2 からは 16 個生成した (計 52 個)。

各試行の時間的スキーマ及び回答方法は、本実験と同じ。ただし 52 回の試行中の各試行で判定の対象とする音としては、旋律 1 から生成された音列に対しては旋律 1 を構成する 9 つすべての音、旋律 2 から生成された音列に対しては $d\sharp, e\sharp, f\sharp, s\sharp$ の 4 つ、計 13 種類の音をそれぞれ 4 回ずつ用いた。予備実験の各試行は、実際には本実験の試行と混合して行われた。

実験の結果、平均律に従う音高を持つ音のみから構成される音列の各音に対し、「ずれている」の回答数が 52 試行中 13 試行 (25%) 以上の被験者が 3 名いた。残り 9 人の被験者が「平均律を許容する」とみなされ、この 9 人の回答データのみが解析処理に用いられた。