

## [招待講演] 音のピッチ知覚

大串健吾

京都市立芸術大学名誉教授

〒610-1102 京都市西京区御陵大枝山町 1-6-10

E-mail: ohgushi.k@view.ocn.ne.jp

**あらまし** 音のピッチ知覚についての科学研究は 19 世紀中頃から活発な実験や議論がおこなわれてきた。ピッチ知覚を主体とする研究を聴覚理論という。ここでは聴覚理論の歴史的な変遷をかんたんに紹介し、さらに複合音のピッチが基本周波数に対応する純音のピッチとわずかにしかも系統的に異なるという実験事実をどのように説明すべきかについて新しい観点から述べる。

**キーワード** ピッチ, 複合音のピッチ, レジデュアピッチ, オクターブ伸長, トーンクロマ

## [Invited Talk] Pitch Perception of Sound

Kengo OHGUSHI

Professor Emeritus, Kyoto City University of Arts

1-6-10 Oheyama-cho Goryo, Nishikyo-ku, Kyoto, 610-1102 Japan

E-mail: Ohgushi.k@view.ocn.ne.jp

**Abstract** Perception of pitch has been studied and discussed for a long time since the middle of the 19<sup>th</sup> century. As times change, hearing theory has changed. In this lecture, changes in the hearing theory are simply introduced. Further, the experimental fact that the pitch of complex tones is slightly but systematically different from the pitch of pure tone corresponding to the fundamental frequency of the complex tone is explained from a new point of view.

**Keywords** Pitch, Pitch of Complex Tone, Residue Pitch, Octave Enlargement Phenomenon, Tone Chroma

### 1. はじめに

音のピッチ (pitch) とは音の高さ (音高) のことである。音の高さといえば単純そうに思われるかもしれないが必ずしもそうではない。一般に楽音 (調波複合音) のピッチはその楽音の基本周波数と同じ周波数の純音 (正弦波音) に等しいとされているが、必ずしもいつもそうではない。またピアノの一つのキー (例えば中央 A 音) に合わせて発声した母音 /イ/ と /ウ/ は基本周波数が同じでも音の高さの強制選択をしてみると、ほとんどの人は /イ/ が高いと判断する。さらに調波複合音の基本周波数に等しい純音のピッチは一般にその複合音のピッチよりも低く

なる。

このように種々の問題を含むピッチ知覚について、さまざまな現象を取り上げて解説する。なお、この解説の内容は、音響学会サイエンスシリーズの「音のピッチ知覚」[1]からピックアップしたものである。また複合音のピッチがその基本周波数と等しい純音のピッチよりも低いという現象を説明する新しい理論を紹介する。

### 2. 聴覚理論の簡単な紹介

#### 2.1 Seebeck と Ohm の論争

Seebeck は、同心円上に等間隔に穴のあいた回転円盤 (サイレン) に対して直角方向から管で空気を吹き付けた。空気流は

穴と穴の間で周期的に遮断され、明確なピッチをもつ音を発生した。また回転数を速めるとピッチが上昇することを示した。さらに穴の間隔を、 $a, b, a, b, \dots$ と交互に異なった値にすると、ピッチは穴の間隔が等間隔で $(a+b)$ であるような円盤のピッチと等しくなった。これらのことから、Seebeck は音響波形の周期という時間情報がピッチの感覚を惹き起こす（時間説）のだと考えた（1841-44）。

これに対して、Ohm は複合音のピッチは音響波形に含まれる基本周波数成分の存在によって生じるのであると考えており、彼らの論争は解決を見ないままに終わった。

## 2.2 Helmholtz の場所説

約 20 年後に Helmholtz は、中耳が非線形特性をもち、これが基本周波数に対応する差音を発生させ、基本周波数に対応するピッチを惹き起こしているのだと考えた。

また Helmholtz は、基底膜は横に張られた多数の線維から成り、その線維の長さは基底膜の幅が異なるので共鳴する周波数が異なり、異なった周波数の音に対しては異なった線維が共鳴して振動し、これが、異なった聴神経を刺激して異なったピッチを発生させるのであると考え、Ohm の主張を支持した。Helmholtz の理論は場所説 (place theory) と呼ばれ、Seebeck の時間説 (temporal theory) に対抗した。その後数十年の間、場所説が支配的であった。Helmholtz のこの説は共鳴説とも呼ばれている。

## 2.3 Schouten のレジデュー理論

Schouten は、差音の発生による場所説に対して否定的なレジデュー理論という画期的な理論を発表した（1938,1940）[2][3]。彼は光学装置とサイレンを組み合わせて周期的パルス列音（調波複合音）を発生する装置を作成した。周期が 5 ms のパルス列音を注意を集中して聴くと、四つの音から成っているように知覚された。すなわち、200 Hz のピッチをもつ鋭い音色の音と、200, 400, 600 Hz の各純音（基音、第 2 倍音、第 3 倍音）を聴き取ることができた。これ

らの各音の聴こえ方は注意の集中の強さに大きく影響を受けた。Schouten はこれらの鋭い音色の音のピッチ（＝基音のピッチ）をレジデューピッチと呼んだ。レジデューは、基音と同じ高さであるが、純音とは違って濁った鋭い音色をもっているため、音色によって区別できたと述べている。

また、基本周波数が 200 Hz でその 9, 10, 11 倍音（1800, 2000, 2200 Hz）から成る調波複合音の各成分周波数を 40 Hz（中心周波数の 2 %）ずつ上昇させると、各成分は 1840, 2040, 2240 Hz となるが、この非調波複合音のレジデューピッチは 204 Hz（もとのレジデューピッチの 2 % 上昇）となった。もしレジデューピッチが差音によって生じるのならば、200 Hz になり、204 Hz にはならないはずである。このことは明確に差音による場所説を否定したものである。

Schouten ら（1962）[4]はレジデューピッチは音響波形の各周期のピーク間の時間間隔という時間情報（fine structure）によって生じると考えた。

## 3. 調波複合音のピッチと基本周波数

上述のように、複合音のピッチはその基本周波数に等しい周波数の純音のピッチと同じであると考えられてきた。ところが、Walliser(1969) [5]は、単極性パルス列を 1.4~2.8 kHz の帯域フィルタを通したときの調波複合音のピッチを調べた。5 人の聴取者が、基本周波数が 230~380 Hz の範囲で比較音である純音とのピッチマッチングを行ったところ、概して 2 % 程度以下ではあるが低くマッチングされていることが見出した。この結果は従来の常識を覆すものである。

また Smoorenburg(1970) [6]は、2 周波複合音（1.8 kHz+2.0 kHz; 2.0 kHz+2.2 kHz）と純音それぞれのピッチを聴取者にマッチングさせたところ、複合音のピッチは基本周波数である 200 Hz よりもわずかに低くなるという結果を得た。

次いで Terhardt(1971) [7]は、基本周波数成分を含む広帯域複合音と純音のピッチマッチング実験を行い、基本周波数が約 800 Hz 以下では純音よりも低くなることを示した。また基本周波

数成分を除去すると、純音とのピッチの差は広がることを見出した。

大串(1976a) [8]は、キャリア周波数が 2000 Hz で変調周波数がそれぞれ 333 Hz, 400 Hz, 500 Hz であるような振幅変調音のピッチを純音を比較音として求めたところ、基本周波数に対応する純音よりは振幅変調音の方が系統的に低いことを見出した。さらに大串(1976b) [9], Ohgushi(1978) [10]は、基本周波数がほぼ 200 Hz から 1.5 kHz の範囲では、広帯域複合音（パルス列音）のピッチは基本周波数に対応する純音のピッチよりも低いことを見出した。このように、複合音のピッチはその基本周波数成分と同じ周波数の純音のピッチよりもわずかではあるが低くなるという多くの実験データが報告されている。

この現象を説明する理論が Terhardt (1974) [11]により提案されている。彼の理論によれば、ピッチはスペクトルピッチ (spectral pitch) とバーチャルピッチ (virtual pitch) に分けられ、純音のピッチや聴覚系で分解可能な複合音の成分音のピッチをスペクトルピッチと呼び、複合音のピッチはスペクトルピッチによって合成 (synthesize) されるのである。またその成分音のピッチを知覚する場合に、隣接音間の相互マスキング (mutual masking) によってその成分音のピッチは周波数の等しい純音のピッチとはわずかに異なるのである。とくに基音のピッチは第 2 倍音以上の成分からのマスキングにより、常に純音単独のピッチよりは低くなるのである。この理論は、いわゆる場所説に立った神経興奮パターンのピークに対応してピッチが決定するというものである。

## 4. 関連する知覚現象

### 4.1 オクターブ伸長現象

二つの純音を継時的に聴き、ちょうど 1 オクターブだけ離れて知覚されるように周波数を調整すると、二音の周波数比は 2 よりもわずか (3 %以下) ではあるが大きくなる。つまり心理的オクターブは物理的オクターブよりも広くなる。この現象をオクターブ伸長現象 (octave

enlargement phenomenon) と呼んでいる。この現象は周波数が高くなるほど顕著になるが、低い方の周波数が 2.7 kHz を越えると聴取者のオクターブマッチングの作業が困難になる (Ward,1954) [12]。この理由は周波数がほぼ 5 kHz を超えるとトーンクロマが消失するからである。オクターブ伸長現象の説明は聴神経の純音刺激に対する 1 周期に対応する発火時間間隔が、周波数が高くなるにしたがって神経細胞の不応期のために系統的に周期よりも長くなっていくという生理実験データ (Rose ら, 1967) [13] を用いて行われている (Ohgushi, 1983) [14]。また Rose らの生理実験データと同じ傾向の生理データが後に McKinney ら(1999) [15]によって報告されている。また 2 オクターブおよびもっと広いオクターブ間の伸長現象は個々のオクターブ伸長の和になる (Terhardt,1983) [16]。

### 4.2 複合音の多重ピッチ

Davis, Silverman & McAuliffe (1951) [17]は、繰り返し周波数 (基本周波数) が 90~150 Hz の周期パルス列を中心周波数が 2 kHz の帯域フィルタに通した調波複合音を作成し、聴取者にこの音を純音とのピッチマッチングを行うように求めた。実験結果によれば、聴取者は基本周波数あるいは 2000 Hz に合わせたが、しばしば 1 オクターブの食い違いがあり、場合によっては 2 オクターブの違いもあった。

Schouten, Ritsma & Cardozo(1962) [4]は、振幅変調音 (変調周波数=200 Hz, キャリア周波数=2000 Hz) のピッチの知覚実験を行った。その結果、ピッチは 200 Hz だけでなく、180 Hz 付近と 220 Hz 付近と複数の調波複合音とマッチングされた。彼らはこの現象を、振幅変調音の隣接する変調周波数周期の微細構造

また大串 (1976a) [8]は、振幅変調音 (3 成分) と純音のピッチマッチング実験において、振幅変調音が調波複合音の場合、基本周波数がほぼ 1 オクターブあるいは 2 オクターブ異なる純音ともしばしばマッチングされた。さらに 5 度だけ離れた純音とマッチングされることもあった。これらの結果は、複合音のピッチの判断は場合

によってはかなりあいまいであり、とくにトーンクロマが等しい場合にはしばしば異なる複数のピッチとして判断されることのあることを示している。

## 5. 2成分音の実験結果と新しい理論

### 5.1 2成分音のピッチ実験

テスト音を2成分複合音とし、純音を比較音としてピッチマッチングを行った4人の聴取者の実験結果（大串，1976b）[9]をまとめて図1に示す。

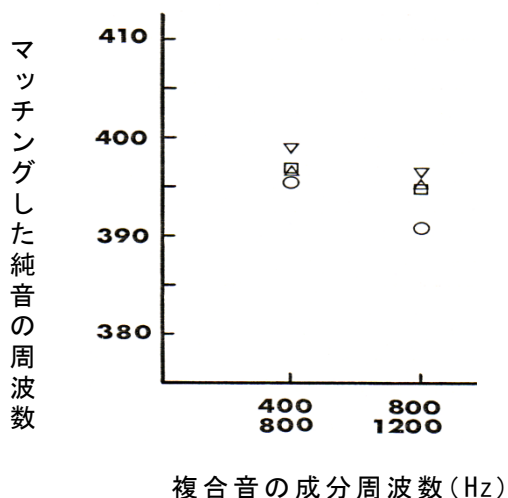


図1 2周波成分複合音のピッチ

図1の結果を要約すると、(1) 400 Hz と 800 Hz の2周波成分音のピッチは400 Hzの純音のピッチよりも低くなる；(2) 基本周波数成分を除去し、第3倍音を加えた800 Hz と 1200 Hz の2周波成分音のピッチはさらに低くなる。これらの結果は、周波数成分がたった2つだけであっても複合音のピッチの低下を示している。

### 5.2 ピッチ知覚のメカニズムに関する理論

上述の結果を説明するために、次のような理論を提案する。すなわち、「純音のピッチを調波複合音のピッチにマッチングする場合、聴取者は複合音の基音成分とのマッチングをするだけでなく、分解できる(resolved)低次倍音とのマッチングを行い、それらの値の間の何らかの加重平均値をとる」。つまり、第2倍音成分との

オクターブマッチングも行われ、また第3および第4倍音あるいはそれ以上の倍音などとのマッチングも行われている可能性が考えられる。聴取者による実験結果の違いは、おそらく倍音の分解能力などの個人差によるもの大きいと思われる。

### 5.3 実験的事実の理論による説明

この理論に従えば、図1の実験的事実は容易に説明できる：

(1) 2成分複合音(400 Hz+800 Hz)のピッチは400 Hzの純音よりわずかに低くなるが、その理由としては、400 Hz付近の純音は、基音の400 Hzとマッチングされるだけでなく、800 Hzの第2倍音とのオクターブマッチングもなされるからである。純音のオクターブマッチングにおいては、心理的オクターブの伸長現象が生じるので、800 Hz部分音の1オクターブ下に感じられる純音の周波数は400 Hzよりも低くなる。したがって、この複合音のピッチは基音とのピッチマッチングおよび第2倍音とのオクターブマッチングとの中間的な値となり、400 Hzよりも低くなるのである。

(2) (800 Hz+1200 Hz)の複合音はさらにピッチが低下する。その理由は、基音成分とのピッチマッチングがなくなり(結合音が生じればピッチマッチングは行われる)オクターブマッチングおよび第3倍音とのマッチングになるからである。

## 6. 討論

Terhardtは場所説により複合音のピッチシフト現象を説明したが、さらに純音のオクターブ伸長現象も同じ原理で説明している。しかし継続的に提示される純音のオクターブ伸長現象はマスキング現象では説明できない。

Petersら(1983)[18]は、基音そのもののピッチはマスキングの影響によってわずかに低下するというTerhardtの理論をチェックするために、部分音(基音を含む)の分析的聴取(analytic listening)の実験を行った。調波複合音の基音成分に注意を集中すれば(完全に集中できれば)

部分音としての基音のピッチは基音の周波数に等しい純音のピッチと等しくなるはずである。

テスト音を、基本周波数を 200 Hz 一定、部分音は 7 次倍音以下の 7 成分音を用いて、聴取者に各部分音のピッチを純音のピッチとマッチングさせた。実験法の一つとしては、テスト用の部分音は最初の 200 ms の間は除去されており、その後の 500 ms の間は加えられた。この途中からの提示により聴取者の注意を引きつけた。実験結果によれば、実験法や提示条件によって本質的には変化はなく、各部分音のピッチは、各成分を単独で提示したときのピッチに非常に近いことが示された。

この結果は Terhardt の理論に対する反論となっており、時間説の方が妥当だと考えられる。

## 文献

- [1] 大串健吾, 音のピッチ知覚 (音響サイエンスシリーズ), 日本音響学会編, コロナ社, 東京, 2016.
- [2] J.F.Schouten, "The perception of subjective Tones," Proceedings of Koninklijke Nederlandsche Akademie van Wetenschappen, Vol.41, No.10, pp.1086-1093, 1938.
- [3] J.F.Schouten, "The Perception of pitch. Philips Technical Review," Vol.5, No.10, pp.286-294, 1940.
- [4] J.F.Schouten, R.J.Ritsma, and B.L.Cardozo, "Pitch of the residue," J. Acoust. Soc. Am., Vol.34, No.8,(2), pp. 1418-1424, 1962.
- [5] K.Walliser, "Zusammenhänge zwischen dem Schallreiz und der Priodenhöhe, Acustica, Vol.21, pp.319-329, 1969.
- [6] G.F.Smoorenburg, "Pitch perception of two-frequency stimuli," J.Acoust.Soc.Am., Vol.48, No.4, pp. 924-942, 1970.
- [7] E.Terhardt, E. "Die Tonhöhe Harmonischer Klänge und das Oktavintervall," Acustica, Vol.24, pp. 126-136, 1971.
- [8] 大串健吾, "複合音の高さの知覚形成のメカニズム," 音響学会誌, Vol.32, No.5, pp. 300-309, 1976a.
- [9] 大串健吾, "複合音の高さの知覚における時間情報の役割," 音響学会誌, Vol.32, No.11, pp.710-719, 1976b.
- [10] K.Ohgushi, "On the role of spatial and temporal cues in the perception of the pitch of complex tones," J.Acoust.Soc.Am., Vol.64, No.3, pp. 764-771, 1978.
- [11] E.Terhardt, "Pitch, consonance, and harmony," J.Acoust.Soc.Am., Vol.55, No.5, pp.1061-1069, 1974
- [12] W.Ward, "Subjective musical pitch," J. Acoust. Soc. Am., Vol.26, No.3, pp 369-380, 1954.
- [13] J.E.Rose, J.F.Brugge, D.J.Anderson, and J.E.Hind, "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol., Vol.30, pp.769-793, 1967.
- [14] K.Ohgushi, "The origin of tonality and a possible explanation of octave enlargement phenomenon," J. Acoust. Soc. Am., Vol.73, No.5, pp.1694-1700, 1983.
- [15] M.F.McKinney and B.Delgutte, "A possible neurophysiological basis of the octave enlargement effect," J. Acoust. Soc. Am., Vol.106, No.5, pp.2679-2692, 1999.
- [16] E.Terhardt, "Auditory universals in music perception," Musicological conference of world music weeks, pp.1-11, 1983.
- [17] H.Davis, S.R.Silverman, and D.R. McAuliffe, "Some observations on pitch and frequency," J. Acoust. Soc. Am., Vol.23, No.1, pp.40-42, 1951.
- [18] R.W.Peters, B.C.J. Moore, and B.R.J. Glasberg, "Pitch of components of complex tones," J. Acoust. Soc.Am., Vol.73, No.3, pp.924-929, 1983.