

聴覚情報処理の概観とピッチ知覚

饗庭絵里子^{†1}

概要: ピッチとは知覚的な「音の高さ」のことであり、音の基本的属性の一つである。ピッチは音波形の周期に応じて変化し、周期が短ければ短いほどより高い音が知覚される。しかしながら、音圧や音色の変化が影響を及ぼすこともあり、周期性だけを手掛かりにピッチ知覚を論じることは難しい。そのため、長年にわたり様々な聴覚理論が提唱され、実験やモデルによる検証が行われてきた。現在でも明確な結論は得られていないものの、心理・生理・数理など様々な分野の成果により、次第に有力な説に絞り込まれつつある。本チュートリアルにおいては、ピッチの定義を確認した後、ピッチ知覚の代表的な理論である「時間説」と「場所説」について、両説に関する生理学的根拠とともに概説する。

キーワード: 場所説, 時間説, 絶対音感

Summary of Auditory Information Processing and Pitch Perception

ERIKO AIBA^{†1}

Abstract: Pitch is the perceptual "height of tone," which is one of fundamental attributes of sound. Pitch is associated with the repetition rate of the waveform of a sound, and higher the repetition rate, the higher is the sound perceived. However, it is difficult to discuss pitch perception using only repetition rate or frequency, because it also changes depending on the sound pressure and timbre. Therefore, various auditory theories have been proposed for many years, and psychological experiments and validation of models have been conducted. Although clear conclusions are still not obtained, credible theories are gradually developing due to the findings of various fields such as psychology, physiology, and mathematics. In the tutorial, the definition of pitch is described, followed by the outline of "temporal theory" and "place theory," which are representative theories of pitch perception. The physiological evidences for both theories are also provided.

Keywords: Place theory, Temporal theory, Absolute pitch

1. はじめに

1.1 ピッチの定義

ピッチとは知覚的な「音の高さ」のことである。American National Standardsによれば「*Pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high.*」と定義されており [1], 日本工業規格 (JIS Z 8106:2000) においても同様に「*聴覚にかかわる音の属性の一つで、低から高に至る尺度上に配列される*」と定義されている [2]. ただし、これらには備考があり「*Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus.*」 [1], つまり「*音の高さは、主として刺激の周波数成分に依存するが、音圧、波形にも関係する*」とされている [2]. この備考は大変短いものであるが、その背景にはピッチを定義することの難しさが隠されている。

例えば、純音に関して、約 4000 Hz 以上の音ではレベルを上げるとピッチが高く聞こえるが、約 2000 Hz 以下の音ではレベルを上げるとピッチは低く聞こえる [3][4]. このような効果は、1-2 kHz の音において約 1% 以下であるが、これより低い/高い周波数の音においては約 5% まで高さが

変化する。また、その変化の大きさや、ずれの方向には個人差も見られる。

複合音の場合、典型的には基本周波数に応じたピッチが知覚される [5][6]. しかし、純音とのピッチマッチング実験の結果は意外にも不安定なものであり、特に調波複合音においては基本周波数ではなく倍音やオクターブ関係にある純音のピッチに等しいとされることも多いようである [7][8][9]. また、最も有名な例外としてはミッシングファンダメンタル現象によって知覚されるレジデュールピッチがあげられる [10][11]. ミッシングファンダメンタル現象とは調波複合音の基本周波数成分を取り除いた場合でも、その基本周波数成分と同じピッチが知覚される現象のことである。具体的には、クリック音を 1 秒間に 200 回の周期で生成した場合、200 Hz の純音の高さに相当するピッチが生じる。この音は、200, 400, 600, 800, ... Hz などの倍音を含んでいるが、このうち 200 Hz の成分を取り除いたとしても、依然として 200 Hz 相当のピッチが知覚される。

以上のように、ピッチは確かに音の高さのことではあるものの、単純に純音の周波数や、複合音の基本周波数によって規定されるものではないと言える。

加えて、ピッチは「*低から高に至る尺度上に配列される*」

^{†1} 電気通信大学
University of Electro-Communications

という一次元的性質だけでは説明できない。これは音楽における音階を思い起こせば容易に理解することができる。音階は、ドから始まり、レ、ミ、ファ、ソ、ラ、シと上行した後、再びドという階名の音にかえるのである[12]。つまり、ピッチには周波数が高くなるにつれて上行するという性質（トーンハイト）に加えて、1 オクターブを周期とする循環的な性質（トーンクロマ）を併せ持っているのである。

ピッチの定義とは、やはり一筋縄ではいかないことがお分かりいただけたと思う。観察された様々なピッチ知覚現象をもとに、多くの研究者が聴覚理論を提唱してきたものの、いまだに明確な結論は得られていない。しかしながら、心理・生理・数理など様々な分野の成果により、次第に有力な説に絞り込まれつつあることも確かである。

次に、特に長く議論の対象となってきたピッチ知覚理論である「時間説」と「場所説」について述べる。

2. 聴覚理論

2.1 時間説と場所説

時間説とは、音に含まれる周期性などの時間情報を手掛かりとしてピッチを知覚しているとする説である。時間説を唱えた Seebeck は、等間隔に穴を開けた回転円盤に管から空気流を吹き付けることでインパルス波を発生させ、それによってピッチを生成した。円盤の回転が速くなればインパルス波が発生する周期は短くなり、それにつれてピッチが高くなる。この実験から、Seebeck はピッチが時間情報によって規定されるという時間説をとった。特に、前述のレジデューピッチは、時間説の強力な証拠とされた。

これに対し、場所説とは、蝸牛基底膜上で共鳴する周波数が場所によって少しずつ異なり（鼓膜に近いほど低い周波数が共鳴）、それによってピッチを知覚しているとする説である。この説は、複合音のピッチはその音に含まれる周波数成分の存在によってのみ知覚されるとする「Ohm の法則」に基づき Helmholtz が提唱した理論である[13]。レジデューピッチが報告されるまでは、場所説が有力であるとされていた。

しかしながら、時間説だけでは説明できない知覚現象が観察されるようになり、新たにパターン認識の原理を適用したモデルが提案されることとなった[14][15] [3]。

2.2 パターン認識モデル

パターン認識モデルは、場所説を踏襲するモデルであり、基本的には二つの段階を含んでいる。第一段階目は、周波数分析器の役割を担っており、粗い帯域フィルタ群によって周波数成分が分解される。この段階が、蝸牛基底膜上の場所に応じて共鳴する周波数が違うという場所説に対応している。第二段階目では、分解された周波数成分を用いてパターン認識が行われる。従って、レジデューピッチのように基本周波数が存在しない音であっても、残された周波

数成分のパターンからピッチを抽出しているとしている。

2.3 生理学的証拠

動物による神経生理学的実験においては、刺激波形の特定の位相に対して生じる神経発火（位相固定）は 4-5 kHz 以上で消失する[16][17]。知覚においても、これ以上の周波数では、純音の周波数弁別する能力は低下し、音楽的な高さの感覚も消失する。従って、4-5 kHz 以下の周波数範囲では時間情報が利用されている可能性が高いと考えられる。また、位相固定が消失する領域であってもピッチが失われるわけではないことから、4-5 kHz 以上の周波数範囲では場所情報が利用されていると考えられる[18]。

また、レジデューピッチを引き起こす高調波成分を提示した場合、基本周波数のみを聞かせた場合と同じように応答するニューロンの存在が一次聴覚野において観察されている[19][20]。つまり、同一のニューロンが時間情報にも場所情報にも反応していることを示している。

3. おわりに

長きにわたり、どちらの説が正しいかについて議論が行われてきたが、現在ではどちらも活用されているという説が有力である。使える手がかりは何でも使う、あるいは状況に応じて最適な情報処理を行っているというのが妥当な見解かも知れない。最近では機械学習を用いた新たな研究も登場しており、この歴史的テーマに大きな展開がもたらされることが期待される。

参考文献

- [1] ANSI, American National Standard: Acoustical Terminology. New York: American National Standard Institute, 1994.
- [2] 日本工業規格 JIS X 8106, 2000.
- [3] E. Terhardt, "Pitch of pure tones: its relation to intensity," in Facts and models in hearing, ed: Springer, 1974, pp. 353-360.
- [4] J. Verschuure and A. Van Meeteren, "The effect of intensity on pitch," Acta Acustica united with Acustica, vol. 32, pp. 33-44, 1975.
- [5] R. W. Peters, B. C. Moore, and B. R. Glasberg, "Pitch of components of complex tones," The Journal of the Acoustical Society of America, vol. 73, pp. 924-929, 1983.
- [6] W. M. Hartmann and S. L. Doty, "On the pitches of the components of a complex tone," The Journal of the Acoustical Society of America, vol. 99, pp. 567-578, 1996.
- [7] L. A. Jeffress, "The pitch of complex tones," The American Journal of Psychology, vol. 53, pp. 240-250, 1940.
- [8] H. Davis, S. Silverman, and D. McAuliffe, "Some observations on pitch and frequency," The Journal of the Acoustical Society of America, vol. 23, pp. 40-42, 1951.
- [9] 大串健吾, "複合音の高さの知覚形成のメカニズム," 日本音響学会誌, vol. 32, pp. 300-309, 1976.
- [10] J. F. Schouten, R. J. Ritsma, and B. L. Cardozo, "Pitch of the Residue," The Journal of the Acoustical Society of America, vol. 34, pp. 1418-1424, 1962.
- [11] 日本音響学会, 音響用語辞典: コロナ社, 2003.
- [12] A. Bachem, "Chroma Fixation at the Ends of the Musical Frequency Scale," The Journal of the Acoustical Society of America, vol. 20, pp. 704-705, 1948.

- [13] 大串健吾, 音のピッチ知覚: コロナ社, 2016.
- [14] J. L. Goldstein, "An optimum processor theory for the central formation of the pitch of complex tones," *The Journal of the Acoustical Society of America*, vol. 54, pp. 1496-1516, 1973.
- [15] F. L. Wightman, "The pattern - transformation model of pitch," *The Journal of the Acoustical Society of America*, vol. 54, pp. 407-416, 1973.
- [16] D. H. Johnson, "The relationship between spike rate and synchrony in responses of auditory - nerve fibers to single tones," *The Journal of the Acoustical Society of America*, vol. 68, pp. 1115-1122, 1980.
- [17] A. Palmer and I. Russell, "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," *Hearing research*, vol. 24, pp. 1-15, 1986.
- [18] B. C. J. Moore, *An Introduction to the Psychology of Hearing*: Emerald, 2012.
- [19] C. Pantev, M. Hoke, B. Lutkenhoner, and K. Lehnertz, "Tonotopic organization of the auditory cortex: pitch versus frequency representation," *Science*, vol. 246, pp. 486-488, 1989.
- [20] H. Riquimaroux and T. Hashikawa, "Units in the primary auditory cortex of the Japanese monkey can demonstrate a conversion of temporal and place pitch in the central auditory system," *Le Journal de Physique IV*, vol. 4, pp. C5-419-C5-425, 1994.