

研究論文

ノッチフィルタが広帯域雑音の音像定位に及ぼす影響

山高 正烈*, 坂本 修一**, 鈴木 陽一**

(2020年10月9日受理)

Effects of Notch Filters on Sound Localization of Broadband Noise

Zhenglie CUI*, Shuichi SAKAMOTO**, Yôiti SUZUKI**

(Received October 9, 2020)

Abstract

The spectral shape of the head-related transfer function plays an important role in sound localization, particularly in the median plane. On the other hand, the timbre of sounds is closely related to their frequency spectrum. Interestingly, humans can separately perceive both the timbre and position of sound sources even though the frequency spectrum of a sound reaching the listener's each ear is a mixture of both frequency spectra. Our previous studies reported that listeners can detect one of the spectral notches in the HRTF, known as N1 and commonly centered around 7 kHz, by the changes it causes on the perceived timbre. In this study, we focus on the spectral notches in the HRTF (N1, N2), of which central frequency may systematically depend on the elevation angle, and analyzed the influence of the notches included in broadband noise on the sound localization perception. Results shows that elevation localization and mutual changes between N1 and N2 could be explained in terms of the excitation pattern.

キーワード: 音空間知覚, 頭部伝達関数, 音像定位, 音色, 知覚手がかり

Keywords : Spatial hearing, HRTF (head-related transfer function), Sound localization, Timbre, Perceptual cue

1. はじめに

人間は音を聞く際に、音の方向のみならず音色も同時に知覚できる。音色とは、音源が何であるかを認知するための手がかりとなる特性であり、音を聞いた主体が音から受ける印象の諸側面の総称と定義される[1]。音色の知覚は、主に音の周波数特性が影響しており、音源の周波数スペクトルによって変化する[2]。一方、耳に到達する音信号には、音源からの直接音だけでなく、反射音や回折音等が含まれており、耳介周りの外耳道や耳甲介腔による共振も生じている[3]。そのため、両耳で観測される音の周波数スペクトルは音源の相対方向に応じて変化する。この周波数スペクトルの変化を音源の位置を変数とする伝達関数として表現したものが頭部伝達関数(HRTF: head-related transfer function [4])であり、人間が音像定位[5]をする際の重要な手がかりとなっている。特に、両耳に入力される音の時間差や強度差がほとんどない正中面や、その変化があまりない矢状面上(両耳からの距

離差が一点の円上, cone of confusion[6])に位置する音源の音像定位については、この頭部伝達関数のスペクトル形状が非常に重要な知覚手がかりとされている(スペクトラルキュー)。このように、人間の耳に入力される音信号は、音色の手がかりと音像定位の手がかりが混在した状態であり、両方とも耳入力信号の周波数特性が手がかりであると考えられている。

ところで、耳入力信号の周波数特性が、音色の知覚と音像定位両方の手がかりとなっている中で、例えば、ある人の声が方向によって音色が変わって聞こえる、若しくは、聞き慣れない音は到来方向が分かりにくいといった現象が起こってもおかしくない。しかし、人間は音色と音像の双方を明確に知覚することが可能であり、音の方向が変わっただけで別の音色に聞こえることを日常生活で経験することはあまりないであろう。即ち、人間は日常的に耳入力信号から音色の情報と音像位置の情報を何らかの仕組みで切り分けて知覚し、それぞれの情報を別々に利用していると考えられる。この機序を解明することは、頭部伝達関数に含ま

* 愛知工科大学工学部情報メディア学科, 〒443-0047 愛知県蒲郡市西迫町馬乗 50-2

Department of Media Informatics, Aichi University of Technology, 50-2 Manori, Nishihisama-cho, Gamagori-shi, Aichi 443-0047, Japan

** 東北大学電気通信研究所, 〒980-8577 仙台市青葉区片平 2-1-1

れる音像定位の手がかり, スペクトラルキューを調べる上で有用な知見になると言える.

しかし, 頭部伝達関数のどのような特徴が正中面音像定位の手がかりとなっているかについては, 未だ断定的な結果が得られていない. 過去の研究では, 2 kHz以上の周波数帯域における頭部伝達関数の概形を再現すれば高精度の音像定位が可能であり, そのスペクトラルキューとして, 仰角によって系統的に変化する複数帯域のレベル差が示唆されている[7]. また, 仰角によって系統的に変化する7 kHz付近に現れる第1ノッチの形状(以下N1)が, 音像定位の工学的制御に有効であるという報告もなされている[6]. しかし, 聴覚系における周波数分析機構をつかさどる聴覚フィルタの特性を考慮すると, N1のように鋭いスペクトル上のノッチは, 元の形状よりもはるかに緩やかで浅い興奮パターンを示すと考えられる[8]. したがって, 音像定位の手がかりはN1そのものの形状やその変化ではなく, 興奮パターンの形状にあると考えるべきである. 実際, 頭部伝達関数の2 kHz以上の帯域を平滑化しても音像定位が可能であるという報告や[7], 8 kHzを中心とする3/4オクターブ帯域においてそれ以外の帯域とは仰角の変化に伴う帯域のレベルの変化傾向が異なる様相を呈することを示し, それらの帯域にスペクトラルキューがあるといった報告もなされている[9].

以上のように, 音像定位の工学的制御にN1が重要であることは明らかであるが, それに起因する聴覚興奮パターン上のスペクトラルキューが何なのかについては統一的な理解が得られていない. これらの背景を踏まえ, 筆者らは音像定位と音色知覚における周波数スペクトルの影響を解明する礎として, N1を模擬するノッチフィルタが広帯域雑音の音色と音像定位に及ぼす影響について検討を進めてきた[9-11]. 我々の先行研究では, 広帯域雑音中に付加したN1ノッチフィルタの帯域幅や中心周波数を系統的に変化させ, それに対応する音色変化の検知限と聴取者自身の頭部伝達関数を比較することで, 音色と音像定位のスペクトラルキューを検討した. その結果, N1の中心周波数の変化は, 音色の変化としては知覚されうるが, 仰角方向の定位方向に充分寄与するとは言えないことが示唆された. 即ち, 広帯域雑音中におけるN1の中心周波数の制御だけでは, 仰角方向の音像定位の手がかりとして不十分であることが明らかとなった.

仰角方向の音像定位の手がかりを探るための先行研究として, 飯田ら[6]は, 聴取者自身の頭部伝達関数と同等の定位精度を再現するためには, 少なくともP1(HRTFに現れる第1ピーク), N1, N2(第2ノッチ)が必要であると報告している. また, 曲谷地ら[9]は, 各

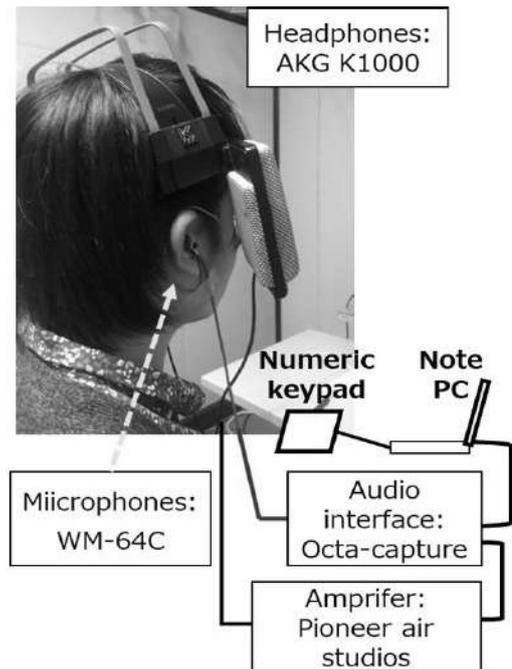


Fig.1 Experimental setup

聴取者の頭部伝達関数の聴覚興奮パターンについて1/4オクターブ帯域間の相互相関を検討し, 8 kHzを中心とする3/4オクターブ帯域間のレベル差を再現すれば音像定位が可能であると報告した. これらの先行研究と, 我々が行ったN1ノッチフィルタと聴覚興奮パターンとの関連についての探索的研究知見[9-11]を踏まえて, 本研究では, N1ノッチフィルタに加えて, 広帯域雑音中にN2を付加することで, N1とN2の互いの変化パターンが音像定位にどのような影響を及ぼすかを検討した.

2. 実験内容

2.1 聴取者と実験環境

正常な聴力を有する, 20代の成人男性4名が聴取実験に参加した.

実験環境の概略を図1に示す. 実験は東北大学電気通信研究所内の防音室で行った. 音の提示には耳介開放型ヘッドフォン(AKG K1000)を用いた. 実験では, 耳入力のスぺクトル形状を一定に保つために, 先行研究[10][11]に習ってヘッドフォンの逆フィルタの更新(詳細は2.2項参照)を行った. そのため, ワイヤを耳介にかける形で測定用小型マイクロフォン(Panasonic WM-64C)を耳珠の真横に設置し, 一連の実験はこのマイクを装着したまま行った.

2.2 逆フィルタ処理

逆フィルタの更新を行うために、まず耳珠横のマイクロフォン位置におけるヘッドフォンの伝達関数(インパルス応答)を、512点の時間引き延ばしパルス(TSP, time-stretched pulse[12])を用いて求めた。次に、ヘッドフォンのインパルス応答の畳込み行列の疑似逆行列を求めた。その際、目的関数には0.2~18 kHzのバンドパスフィルタを用いた[13]。したがって、この逆フィルタに所望の周波数特性を持つ広帯域音を通しそれをヘッドフォンに入力すると、出力音は耳珠横のマイクロフォン位置において入力と同じ周波数スペクトルを示すことが期待できる。

逆フィルタの処理に要する時間は3秒程度であり、セッション開始直前と10試行終了ごとに行った。聴取者の体勢が安定している状態で録音を行うために、10試行ごとの逆フィルタ算出時はヘッドフォンからアナウンスを流し聴取者に通知した。その後、聴取者が任意のタイミングでキーボードを押すとTSPの録音出力と逆フィルタの更新処理が行われた。以上の一連の流れを実験前の練習セッションでも行い、聴取者に手順の確認を求めた。

2.3 実験刺激

実験刺激は、ピンクノイズ(帯域0.2~18 kHz)を基に、それにP1とN1, N2を付加して生成した。P1の中心周波数は4.0 kHzで固定し、N1は、中心周波数が6.0 kHzから+0.5 Cam^{注)}ずつ8.0 kHzまで上昇する正方向へのシフトパターン(正シフト)を設けた。一方N2は、中心周波数が11 kHzから-0.25 Camずつ10kHzまで下降する負の方向へのシフトパターン(負シフト)を設けた。このように、N1の中心周波数の上昇に伴いN2の中心周波数が下降する、計5種類の刺激を作成した。

以下、N1ノッチフィルタを例に、フィルタの設計について説明する。設計に当たっては、 g , r , ω_0 をパラメータとした式(1)で表される2次のIIRフィルタを用いて反共振ノッチフィルタを作成した。ここで、 f_n , G_c , G_n , B は、それぞれ中心周波数[Hz]、帯域幅を定めるカットオフのレベル[dB]、ノッチの深さ[dB]、 G_c における帯域幅[Hz]を表す。なお、 f_s はサンプリング周波数である。P1, N1, N2とも深さ(G_n)は-20 dB、帯域幅はノッチ両側の平坦な帯域のレベルから-3 dB低下する周波数間の値であり、+1.5 ERB^{注)}に固定した。

$$H(z) = \frac{1-2rcos\omega_0z^{-1}+r^2z^{-2}}{1-2\frac{g}{r}cos\omega_0z^{-1}+\frac{2g-1}{r^2}z^{-2}} \quad (1)$$

$$g = \frac{1}{1+tan\left(\frac{B_0}{2}\right)\frac{\sqrt{1-A_c^2}}{A_c}}$$

$$r = 1 - A_n \frac{B_0}{\pi}, \quad \omega_0 = 2\pi \frac{f_n}{f_s}$$

$$A_c = 10^{\frac{-G_c}{20}}, \quad A_n = \frac{1}{A_c} 10^{\frac{-G_n}{20}}, \quad B_0 = 2\pi \frac{B}{f_s}$$

2.4 実験手続き

N1とN2ともに中心周波数のシフト量が0 Camの場合(基準中心周波数N1=6.0 kHz, N2=11 kHz)を参照音Rとし、その他4種類の刺激音を比較音Xとした。参照音Rと比較音Xをランダムな順序でヘッドフォンから提示し、1回目の刺激音に対して2回目の刺激音が仰角方向の上方が下方のどちらに偏移したかを回答するよう求めた(2肢強制選択)。参照音と比較音の提示時間はともに0.5秒であった。

実験においては、参照音Rと比較音Xの刺激対4種類をランダムな順番で15回ずつ提示した。これを1セッションとし、各セッションの間に休憩をはさみ、4セッション(刺激対4種類×15回×4セッション=240試行)実施した。

3. 実験結果

図2に、各聴取者のN1とN2の中心周波数偏移に対する仰角定位の弁別能を示す。弁別能は $d' = 1$ となる周波数偏移とした。ここで、 d' [14]の算出に当たっては、4セッションの実験結果(音像の仰角定位方向の偏移)に対し、N1の中心周波数が高くなるにつれて(N2の中心周波数が低くなるにつれて)音像が上方に定位するという仮定の下で正答率を求め、その値を元に導出した。グラフの横軸はN1の中心周波数の偏移量を $ERB_{N_{number}}$ スケールで示した値であり、単位はCamである。グラフの縦軸は d' である。グラフ中の●点は各中心周波数のシフトパターンにおける実験結果であり、実線は最小二乗回帰直線である。

図2で弁別限($d' = 1$ となる時の中心周波数の偏移)は、実線が破線と交わる場合のX軸の値となる。図2から、聴取者4名のうち1名のみ(Listener4) d' が1を超えており、その弁別限は1.43 Camであった。ただしListener1の d' の値も、N1の中心周波数の正の方向へのシフトに従い上昇する傾向が見て取れる。即ち、

注) 聴覚フィルタは中心周波数に応じて帯域幅が変化し、その値はERB (equivalent rectangular bandwidth)として表現される。中心周波数によって幅を変えるフィルタが等間隔に並ぶように変形した周波数軸を $ERB_{N_{number}}$ スケールと呼び、単位はCamと表される。

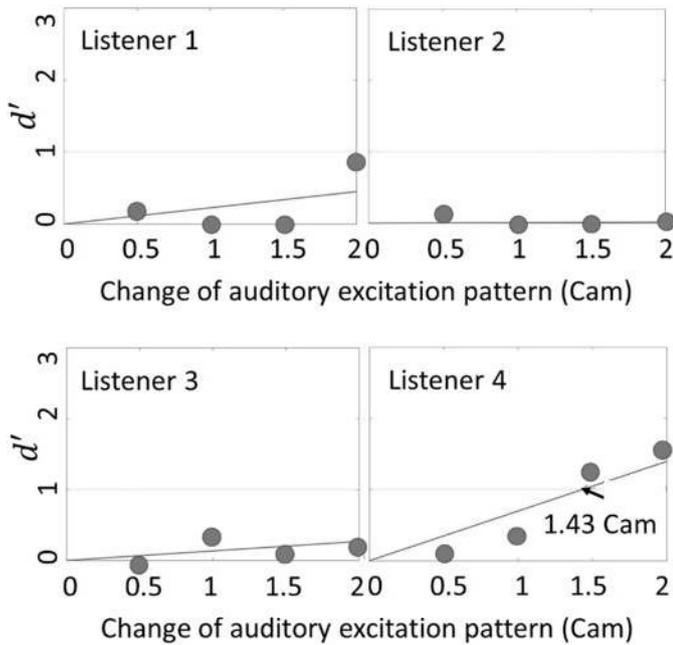


Fig.2 Timbre discrimination performance (d') as a function of the change of estimated auditory excitation pattern caused by the center frequency deviation of N1

N1とN2の中心周波数の偏移が、仰角方向の音像位置の弁別には有効に働くグループと、それほど有効に働かないグループに分かれることが示された。

4. 考察

聴取者 4 名のうち 2 名のみではあれ、このグループについてはN1 中心周波数が正の方向へシフトし、合わせてN2の中心周波数が負の方向へシフトするという周波数偏移パターンを用いた場合、 d' の値の上昇傾向が示された。これから、N1 とN2 の中心周波数の遷移が仰角方向の音像定位に何らかの影響を及ぼす可能性が考えられる。その原因を検討するために、聴取者自身の頭部伝達関数を用いて、実際のN1、N2の中心周波数が、頭部伝達関数上では仰角に応じてどのように変化するかを調べた。

図 3 に、各聴取者の正中面頭部伝達関数における、音源の仰角方向の角度とN1、N2 の中心周波数の関係を示す。●はN1、▲はN2、実線と破線はそれぞれの最小二乗回帰直線を示す。図 3 から、N1 とN2 の関係性は、図 2 の実験結果と同様に 2 つのパターンに分けられることが分かる。即ち、Listener2 やListener3 に見られるように、仰角の上昇に伴ってN1 とN2 の中心周波数がともに同程度高くなる場合と、Listener1 とListener4 のように、仰角の上昇に伴ってN1 の中心周波

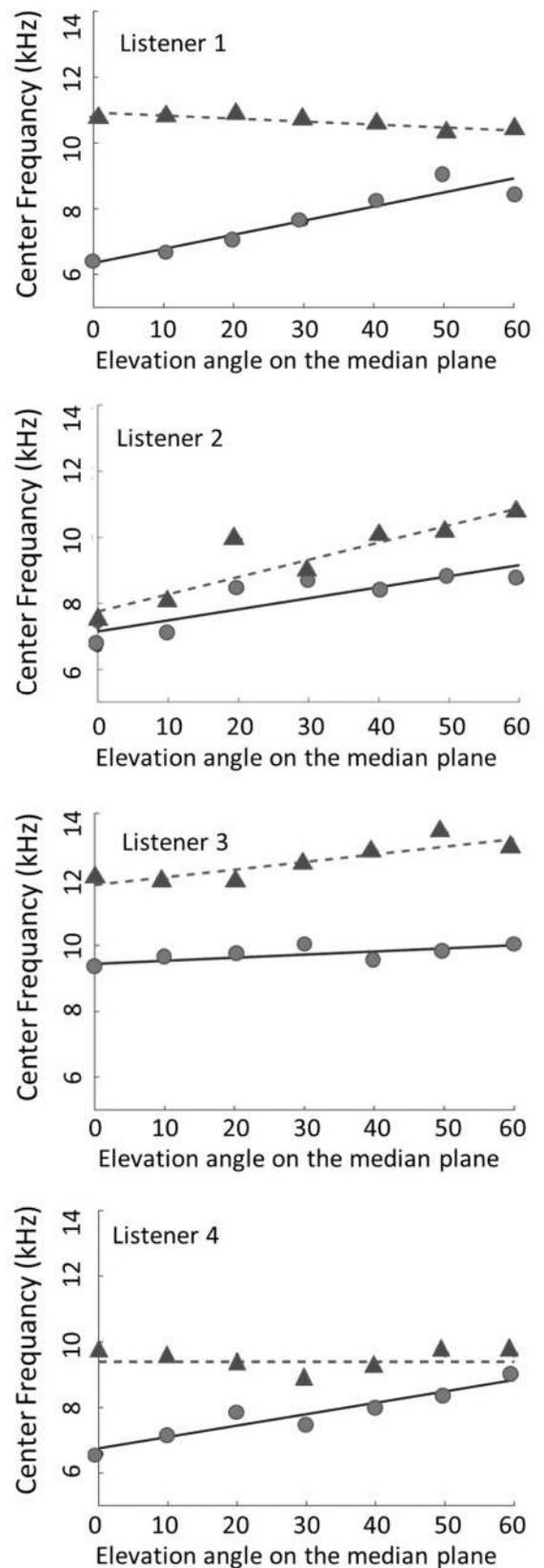


Fig.3 N1 and N2 center frequencies in HRTF as a function of the elevation angle of the sound source (●:N1, ▲:N2)

数は高く、N2 の中心周波数は低くなる若しくは変わらない場合である。便宜的に、前者のN2 の中心周波数が仰角に応じて高くなるパターンを「正シフト」、低くなることを「負シフト」と呼ぶことにする。

図2の結果と、図3における各聴取者ごとの頭部伝達関数の中心周波数の変化パターンを照らし合わせて見ると、N1、N2の中心周波数のシフト量に比例して音像定位の弁別能の向上が見られたListener1とListener4は、「負シフト」パターンに属していた。即ち、聴取者自身の頭部伝達関数においては、N1とN2の中心周波数が互いに逆方向に偏移しているパターンであることが分かる。他方、仰角方向の音像定位の変化が見られなかった聴取者らは、「正シフト」パターンに属していた。即ち、聴取者の頭部伝達関数において、N1とN2がいずれも仰角の上昇に対して中心周波数が高い側に偏移しているパターンであった。

本実験で用いた刺激は、「負シフト」のノッチ中心周波数の偏移を模擬したものであり、N1とN2の偏移の仕方と、自身の頭部伝達関数のN1、N2の偏移の仕方が類似した聴取者のみにおいて、仰角の弁別が可能であったことが考えられる。このことから、仰角の弁別においては、N1とN2の互いの偏移の様相がキューとなる可能性が考えられる。

また、本実験で用いたN1は、中心周波数6 kHzを始点として0.50～2.0 Camの偏移、N2は、中心周波数11 kHzを始点として、0.25～-1.0 Camの偏移であった。ところで、ノッチフィルタの「負シフト」パターンに属する聴取者自身の頭部伝達関数において、N1やN2は必ずしもこの範囲に収まっているとは限らない。しかし、そのような聴取者であっても、ノッチの偏移の様相が似通っていれば仰角方向音像位置の弁別の可能性が示されたことから、仰角定位方向の変化は、ノッチのある程度の広さを持った帯域における興奮パターンの変化で説明可能であることが考えられる。

一方、N1ノッチフィルタの中心周波数のシフト量が音色弁別限に及ぼす影響を調べた先行研究[11]によると、N1の中心周波数偏移に関する弁別限が聴覚興奮パターン上でおよそ13.7 dB・Camのレベル差であることが示されている。そこで、本実験におけるノッチフィルタの「負シフト」に属する聴取者Listener1とListener4の聴覚興奮パターン上のレベル差を求め、音像定位の弁別限と音色弁別限の聴覚興奮パターンを比較した(図4参照)。

図4の「負シフト」パターンに属する聴取者らの弁別限における聴覚興奮パターン上のレベル差に着目すると、Listener1は25.8 dB・Cam、Listener4は16.2 dB・Camであった。即ち、仰角方向の音像位置を弁別し

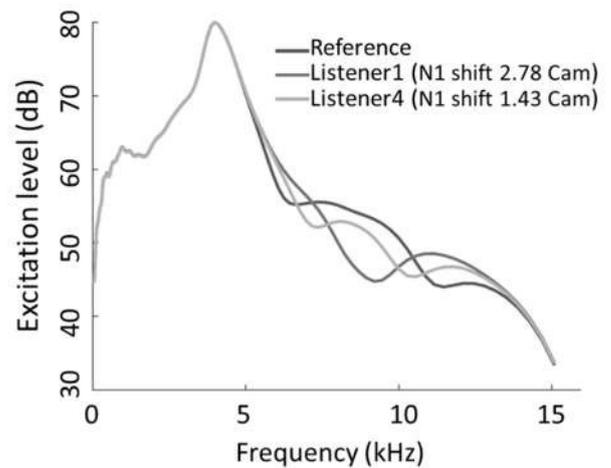


Fig.4 Estimated auditory excitation patterns (Cam)

得るレベル差が、先行研究で示した音色の変化を知覚し得るレベル差よりも大きい値を示していることが分かる。即ち、今回の実験における「負シフト」パターンの聴取者らは、音像定位方向の変化と同時に音色の変化も同様に知覚し得る可能性があったと考えられる。ただし、本研究では音像定位と音色弁別の相互関係に関しては検討するに至っておらず、今後更なる検討が必要である。

以上から、頭部伝達関数に現れるノッチやピークは聴覚興奮パターン上において知覚可能な程度の変化を生じさせ、同時にそれらの相互的な変化の様相が音像定位手がかりとなっている可能性が示された。今後は、仰角の上昇に伴って、N1とN2の中心周波数がともに同程度高くなる「正シフト」パターンを用いて同様の実験を行い、N1とN2の互いの偏移の様相が仰角方向の音像定位に及ぼす影響について、更なる検討を進める予定である。

5. おわりに

本研究では、頭部伝達関数の周波数スペクトルに典型的に現れるノッチやピークに着目し、それらが正中面音像定位に及ぼす影響を聴覚興奮パターンに基づいて検討した。具体的には、頭部伝達関数上に典型的に現れるノッチの音特性知覚への寄与を、N1、N2の中心周波数に着目しながら仰角の弁別実験を行った。その結果、仰角方向の音像定位は、N1とN2の相互変化のパターンによって形成される聴覚興奮パターンの変化の様相によって説明できる可能性が示された。

しかしながら、N1とN2の相互変化によって仰角定位方向に偏移が生じる際に、音色知覚がどのように変化

するのかという、音色と音像定位の相互関係については不明なままである。今後、本研究の知見を基に、音像定位と音色弁別の知覚機序の更なる解明とそれらの相互作用という観点からパラメータを増やして更なる検討を進める計画である。

謝辞

本研究の一部は、JSPS科研費19H02396の助成を受けた。本研究の聴取実験にご協力いただいた東北大学大学院情報科学研究科の日吉啓、及川隼平両氏に感謝する。

参考文献

- [1] 日本音響学会, “音響用語辞典,” コロナ社, 東京, 2003.
- [2] 小澤賢司, “聴覚系にける音色知覚過程に関する研究,” 東北大学審査博士論文, 1993.
- [3] E. A. G. Shaw, “Transformation of sound pressure level from the free field to the eardrum in the horizontal plane,” *J. Acoust. Soc. Am.*, vol.56, no.6, pp.1848-1861.
- [4] M. Morimoto and Y. Ando, “On the simulation of sound localization,” *J. Acoust. Soc. Jpn. (E)*, vol.1, pp.167-174, 1980.
- [5] J. Blauert, “Spatial Hearing- The Psychophysics of Human Sound Localization,” Revised edition (MIT Press, Cambridge, Mass.), 1996.
- [6] K. Iida, M. Itoh, A. Itagaki and M. Morimoto, “Median plane localization using a parametric model of the head-related transfer function based on spectral cues,” *Applied Acoustics*, no.68, pp.835-850, 2007.
- [7] F. Asano, Y. Suzuki and T. Sone, “Role of spectral cues in median plane localization,” *J. Acoust. Soc. Am.*, no.88, pp.159-168, 1990.
- [8] 北島周, 鈴木陽一, 坂本修一, “周波数振幅スペクトル上のノッチによる音色変化の検出限と弁別限,” 日本音響学会講演論文集, 2-P-9, pp.819-820, 2015.
- [9] 曲谷地哲, 岩谷幸雄, 鈴木陽一, “仰角依存性の強い帯域のレベルを模擬した頭部伝達関数による正中面定位,” 聴覚研究会資料, vol.42, no.1, pp.47-52, 2012.
- [10] 齋藤優一, 鈴木陽一, 坂本修一, “ノッチフィルタが広帯域雑音の音色弁別に及ぼす影響,” 日本音響学会聴覚研究会資料, pp.77-82, 2016.
- [11] 日吉啓, 鈴木陽一, 坂本修一, “広帯域ノイズ中のノッチ中心周波数の偏位が音色と音像定位に与える影響,” 電子情報通信学会技術研究報告, EA2018-44, pp.103-108, 2018.
- [12] Y. Suzuki, F. Asano, H. Y. Kim, T. Sone, “An optimum computer-generated pulse signal suitable for the measurement of very long impulse response,” *J. Acoust. Soc. Am.* vol.97, no.2. pp.1119-1123, 1995.
- [13] 川浦純一, 鈴木陽一, 曾根敏夫, 相馬次郎, “悪条件の伝達関数の逆特性推定のための新しい方法,” 電子通信学会論文誌, J69-A, pp.914-917, 1986.
- [14] G.A. Gescheider, “Psychophysics: the fundamentals,” Psychology Press, 2013.