

## 帯域雑音エッジ聴音程の非線形特性

陳施佳 李珍咏 伊東乾

### Non-Linear Characteristics of Band Noise Edge Listening

Shijia CHEN, Jinyoung LEE and Ken ITO

#### 梗概

高い  $Q$  値をもつデジタルフィルタで狭帯域雑音を系統発生させ、上界が可聴域上限を超えるようにした、このような刺激列と、対応する正弦波の刺激列を用い、上行と下行の双方で音程識別の認知テストを行ったところ、強いヒステリシスが観測され、帯域雑音での高域周波数認知は、正弦波と比較して弁別が著しく困難であることが判明した。音声言語の子音聴など、ヒト聴覚の高帯域での認知は大半が帯域雑音であり、旧来の正弦波を用いる聴覚検査では検出できなかった、ヒト高周波聴の非線形な特性の存在が明らかになった。

#### Abstract

Band noise stimuli with steep spectral rise-up, whose high frequencies boundaries are beyond human audible range, are generated by use of high  $Q$ -value digital filters. By use of such series of band noises and sinusoidal stimuli pitch discern tests are performed.

Extremely strong hysteresis is observed in the listening of up-ward and down-ward frequency recognition. In the band noise test, pitch discern is considerably more difficult than in the series of sinusoidal stimuli. Most human high-frequency listening has much to do with consonants in spoken language. Our results show new non-linear characteristics of human high-frequency hearing that could tell difficulties in spoken language distinction which were unclear with those conventional hearing tests using sinusoidal stimuli sound.

Key words: band noise, high frequency, edge listening, Fermi frequency, nonlinear hysteresis

\*) 東京大学大学院情報学環 作曲=指揮・AI生命倫理/情報動力学研究室

Division of Composition/Conducting, Information Statistic Dynamics and AI Life ethics,

Interfaculty Initiative in Informatics, The University of Tokyo

7-3-1 Hongo Bunkyo-ku, 113-0032 Tokyo JAPAN

## 1. はじめに

ヒト内耳の蝸牛管では、可聴域の上限 20,000Hz 付近の反応領域は螺旋状の管の入口近辺に位置し、また下限 20 Hz 付近の反応領域は最深部に位置するとされる [Fig.1]。蝸牛管は音波を周波数ごとにアナログ分波し、対応するニューラルなシグナルを中枢神経系に伝達する。このようにして、ヒトは音を認識している。

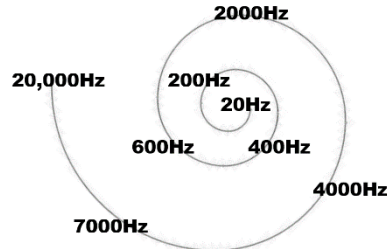


Fig. 1 Frequency recognition in cochlea

ハンガリーの聴覚生理学者 G. von Békésy は、蝸牛の微細構造と機能を解明し、1962 年にノーベル医学生理学賞を授与された。ノーベル賞受賞の翌年 von Békésy は人間のエッジ聴の現象を初めて報告した。これは帯域雑音の聴取で端（エッジ）付近の音高のみが認識されるという特異的な非線形認知現象である。現実には人間が耳にする音はすべて複数の周波数成分を有する帯域雑音である。オシレーターで人工的に生成した、自然界には存在しない正弦波の純音信号であっても、実在するスピーカーから音波が放射される段階で、不可避免的に一定の線幅が付随してしまう。

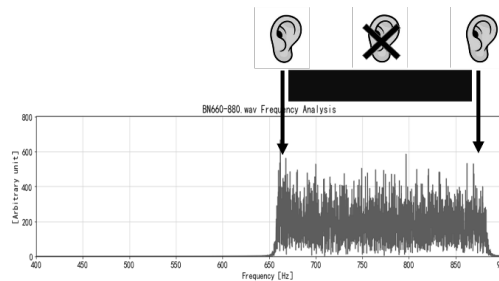


Fig. 2 Edge listening of band noises

Fig.2 に示すように、人間が帯域雑音を聴取するとき、蝸牛など聴覚末梢系では帯域雑音に含まれるすべての周波数成分に反応するが、中枢神経系、即ち脳や私たちの主観的な意識は帯域雑音の上下両端付近の「エッジ音高」しか認識しない。

エッジ聴の進化史的な出現は、人類より遥かに古いと考えられる。生物、特に水中から上陸した脊椎動物は、雑音の多い環境から必要な音のみを選択し、効率的に認知演算することで、個体並びに種の生存可能性を大きくしてきたと思われるからである。

人間が言語音声を目にしてそこに言語の意味と情動を感じるのは、狭域に集中した帯域雑音のなかから、何等かの成分を取り出すエッジ聴のメカニズムに依っている。

人間の声を構成する音素は、主に低周波の母音と高周波の子音から成立しており、単語を構成する母音と子音から音声形成され、通常のコミュニケーションが可能となる。[1]

低周波成分を欠き、連続スペクトルを含む高周波成分だけで構成される囁き声であっても、大半の意味を伝達することができる。「k」「t」「sh」の場合、聞き取りにくいことが多い。

実際に人間が話す音声言語の周波数は 100Hz~10,000Hz 程度の範囲内に収まる。他方、人間の可聴域は 20 Hz~20,000Hz 程度である。ここで、従来の正弦波を用いる聴覚検査か考えると、多くの人異なる子音を明確に識別できるはずである。しかし、現実に聴取される子音は帯域雑音であり、以下に示すように正弦波の聴取と帯域雑音のそれとは大きく異なる生理的な特徴を持つことを、今回の我々の測定は強く示唆している。

## 2. エッジ聴を用いた認知評価実験

### 2.1 認知評価実験の設定

エッジ聴聴取の実験に用いられる帯域雑音サンプルは、von Békésy 以来の慣習で Fig.4-a のような  $(0, f_{edge}]$  あるいは  $[f_{edge}, \infty)$  の开区間エッジサンプルを使用していた。我々はこれを改めデジタル・ベースの高いQ値を持つバンドパス・フィルタにより、Fig.4-b のような両端の周波数が制約された閉区間の帯域雑音サンプル  $[f_{edge1}, f_{edge2}]$  を初めて使用した。

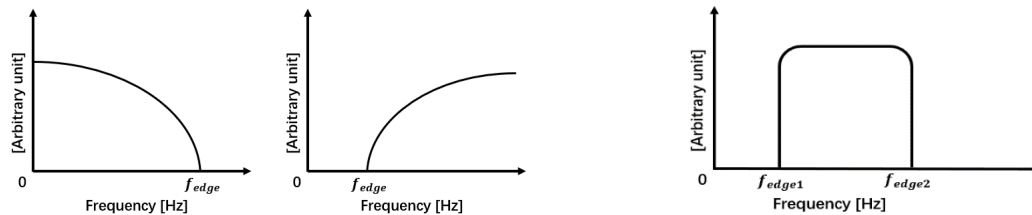


Fig. 4-a Conventional edge listening sample

Fig. 4-b Our novel edge listening sample

帯域雑音サンプルのフィルタ両端周波数は純正調に基づく半音階から計算して与えた。まず、2,500Hz から始まる正弦波の刺激列サンプルを、純正調に基づく半音階で作成した。

次に、正弦波刺激の周波数を下端として帯域雑音刺激列を構成するべく、少しずつ幅を広げた。バンド幅は短2度 ( $\frac{17}{16}$  倍の周波数までの連続スペクトル)、長2度 (同様に  $\frac{9}{8}$  倍の周波数まで)、長3度 ( $\frac{5}{4}$  倍まで)、完全5度 ( $\frac{3}{2}$  倍まで)、オクターブ上 (2倍まで)、長10度うえ ( $\frac{5}{2}$  倍まで) と徐々に広げた。このように系統作成した帯域雑音半音階刺激サンプルの周波数構成を Fig.5 に一部示した。

正弦波	短2度上	長2度上	長3度上	完全5度上	オクターブ上	10度上
2500 Hz	2500 ~2656 Hz	2500 ~2812.5 Hz	2500 ~3125 Hz	2500 ~3750 Hz	2500 ~5000 Hz	2500 ~6250 Hz
2656 Hz	2656 ~2822 Hz	2656 ~2988 Hz	2656 ~3320 Hz	2656 ~3984 Hz	2656 ~5313 Hz	2656 ~6641 Hz
2812.5 Hz	2812.5 ~2988 Hz	2812.5 ~3164.1 Hz	2812.5 ~3515.6 Hz	2812.5 ~4218.8 Hz	2812.5 ~5625 Hz	2812.5 ~7031.3 Hz
3000 Hz	3000 ~3187.5 Hz	3000 ~3375 Hz	3000 ~3750 Hz	3000 ~4500 Hz	3000 ~6000 Hz	3000 ~7500 Hz
⋮	⋮	⋮	⋮	⋮	⋮	⋮
16666.7 Hz	16666.7 ~17708 Hz	16666.7 ~18750 Hz	16666.7 ~20833.3 Hz	16666.7 ~25000 Hz	16666.7 ~33333.3 Hz	16666.7 ~41666.7 Hz
17777.8 Hz	17777.8 ~18889 Hz	17777.8 ~20000 Hz	17777.8 ~22222.2 Hz	17777.8 ~26666.7 Hz	17777.8 ~35555.6 Hz	17777.8 ~44444.4 Hz
18750 Hz	18750 ~19922 Hz	18750 ~21093.8 Hz	18750 ~23437.5 Hz	18750 ~28125 Hz	18750 ~37500 Hz	18750 ~46875 Hz
20000 Hz	20000 ~21250 Hz	20000 ~22500 Hz	20000 ~25000 Hz	20000 ~30000 Hz	20000 ~40000 Hz	20000 ~50000 Hz

Fig. 5 Boundary frequencies of edge listening band-noise-stimuli

帯域雑音サンプル両端の値を決定した後、高い Q 値をもつデジタルフィルタを用い、ホワイトノイズをフィルタリングして、帯域雑音サンプルを作成した。Fig.6 はホワイトノイズにフィルタを通した 2500Hz から 2656Hz の短 2 度幅帯域雑音サンプルのスペクトルである。

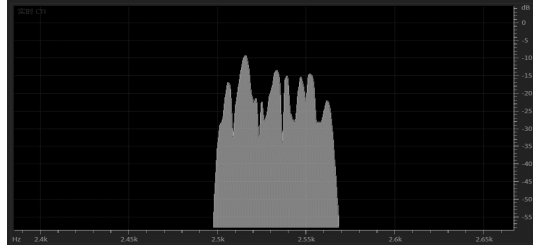


Fig. 6 Spectrum of minor-second-width band-noise from 2500 Hz to 2656 Hz

このように作成したサンプルを用い、人間の可聴域よりも高い周波数領域まで半音階的に上昇する刺激列を系統作成した。このサンプルをまず上行して提示、続いて最高周波数帯域から下降するように提示し、被験者の聴取をしらべた。

10代から20代の被験者7名に目を閉じた状態で刺激を聴覚提示し、聴取不可能となった時点を記録、また再び音が聴こえ始めた時点を記録し、結果をまとめた。

## 2.2 実験結果と考察

Fig.7 に正弦波半音階刺激列による聴取実験の測定結果を示す。表中の「○」は聴取可能、「×」は聴取不可能となる周波数を示す。

被験者 D と被験者 E の回答に非対称性がある以外、5 人の被験者が 2,500 Hz からの半音階上昇中に 16,000 Hz までの正弦波が聞こえ、16,666 Hz 以上の正弦波は聴取できなかったことが確認できる。同時に、20,000Hz からの下降時には、16,666Hz までの正弦波が聞こえず、16,000Hz からは再び聞こえるようになる。即ち、正弦波刺激源では、半数以上の被験者で、上向下向共に同一の可聴域上限閾値が観測される「対称的」な結果が得られた。

20000												
18750												
17500												
16666												
16000	○	×	×	×	○	○	×	×	×	○	○	
⋮												
2500	○	×	×	×	○	○	×	×	×	○	○	
周波数 [Hz]	上向	下向	上向	下向	上向	下向	上向	下向	上向	下向	上向	下向
	被験者A	被験者B	被験者C	被験者D	被験者E	被験者F	被験者G					

Fig. 7 Less Hysteresis with Sinusoidal Semitone Stimuli series

次に、短二度幅の帯域雑音半音階刺激源での応答を調べた [Fig.8]。これは、この程度の帯域幅より狭くなると、エッジ音高が上下二つに弁別困難になるため、この帯域幅を選んだ。短二度幅のサンプルを用いた試行では、上行下行の間に強いヒステリシスが観測された。

21250												
21000												
20000												
19000												
18750												
18000												
17777.8												
17000												
16666.7												
16000	○	×	×	×	○	○	×	×	×	○	○	
15938												
15000	○	×	×	×	○	○	×	×	×	○	○	
周波数 [Hz]	上向	下向	上向	下向	上向	下向	上向	下向	上向	下向	上向	下向
	被験者A	被験者B	被験者C	被験者D	被験者E	被験者F	被験者G					

Fig. 8 Much Hysteresis with minor-second-width Band-noise Stimuli series

Fig.8 に示すように、すべての被験者で上向では聴取不可能だった高周波帯域が下向では聴取されている。短2度幅の帯域雑音は、現実には乱数で周波数と振幅を振り分けられた音波から構成されており、言わば周波数の「ゆらぎ」だけで構成され（超狭域雑音）、認知的に側抑制される「不感帯域」が存在しない。このような帯域雑音では、上端付近の「エッジ音高」と下端付近の「エッジ音高」を別の音程として聞き分けることができず、二つの音程の間に揺動する単一音程が認識される。ゆらぎを含む刺激源の提示により、Fig.7 にしめした正弦波半音階刺激源の測定では聴取不可能だった高周波帯域の音が認知されたことになる。

更に帯域幅を長2度、長3度と拡大すると、半数以上の被験者で上下向における対称性が回復してゆき、帯域幅を完全5度に拡大すると、Fig.9 に示すように、全ての被験者で上下向対称性が回復した。

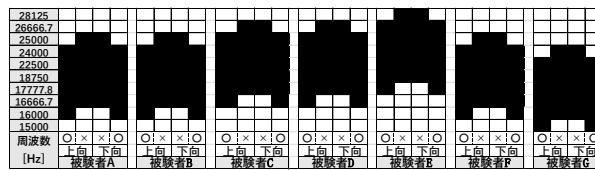


Fig. 9 Less Hysteresis with perfect-fifth-width Band-noise Stimuli series

ところが、帯域幅を1オクターブに拡大すると、Fig.10 に示すように、異なる傾向が乱立する状況が観測された。更に帯域幅をオクターブ以上、即ち長10度に拡大すると、再び全ての被験者で上下向対称性が回復した。

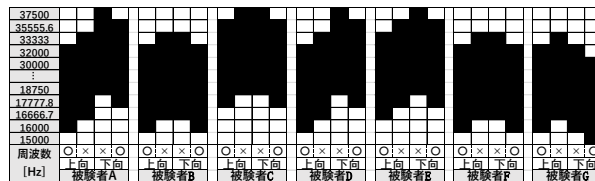


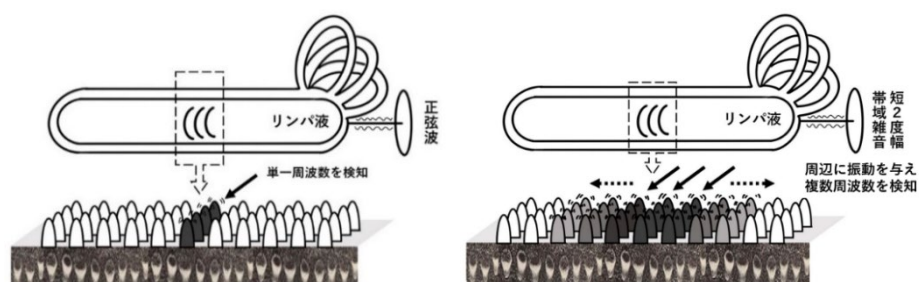
Fig. 10 Result of octave-width Band-noise Stimuli series

オクターブ幅の帯域雑音サンプルについては複雑な認知メカニズムが考えられるため、より詳細な後続実験を準備している。

### 3. まとめと展望

ヒト内耳の蝸牛管はリンパ液で満たされており、鼓膜からの振動が伝わると液相が振動、蝸牛管内基底膜上にある有毛細胞群がそれらを検知し、神経発火のインパルスを発生させる。聴覚刺激のインパルス群は蝸牛神経を通じて中枢神経系にもたらされ、私たちの意識はそれを認識する。

単一の周波数を持つ正弦波が蝸牛管内に導かれると、リンパの液相は狭い範囲で振動し、有毛細胞群が検知する周波数も限局されたものになると思われる[Fig.11]。これに対し、短2度幅で周波数成分に広がりを持つ帯域雑音では、リンパの液相は本来検知されるべき帯域幅を超えて、周辺の有毛細胞群に振動を与え、正弦波の場合とは異なる神経インパルスを発生させる可能性がある[Fig.12]。このためヒステリシスが発生していることが考えられ、後続実験での検証を検討している。



Left: Fig. 11 Schematic view of cochlea-hair cells-excitation after listening sinusoidal stimuli.

Right: Fig. 12 Schematic view of cochlea-hair cells-excitation after listening minor-second-width band-noise stimuli.

ただし、長2度を超える広帯域幅の雑音でヒステリシスの効果が消失する現象、並びにオクターブでの複雑な応答については、これだけでは説明がつかない。副次的な可能性として、側抑制効果とのトレードオフが考えられるだろう。もっぱら狭帯域での「ゆらぎ成分」だけで構成され、側抑制の効果がみられない「(超)狭帯域雑音」では上行と下行で認知特性に変化が見られる。これに対し、側抑制の効果によってバンド幅の上下端に二つの音程が聴取され、中間の音程が聞き取られない広帯域雑音では、振動の周波数方向でのゆらぎの伝播も抑制される可能性があるため、これを確認する追加実験を準備している。

**結論** 高い  $Q$  値を持つデジタルフィルタを用いて帯域雑音刺激列を作成、対応する正弦波刺激列を併用して被験者の聴取実験を行った。その結果、短2度の狭帯域雑音幅では、正弦波では観測されない周波数上下行時の強いヒステリシスが観測された。帯域幅を広げると再び正弦波同様の対称性が回復するが、オクターブ幅の帯域雑音刺激では、複雑な応答が見られた。さらに、帯域幅を広げると、再び対称性が回復した。従来の、正弦波を用いた聴覚検査では、人間の音声言語の聴取、とりわけ子音の弁別など、会話の現実に即した特徴量を検出できていない可能性が考えられる。今回の測定では、狭帯域雑音に特異的なヒト高周波聴の非線形特性を初めて見出すことができた。今後、補聴器の開発など、福祉目的の応用につながる基礎的な研究を、さらに推進してゆく予定である。

#### 参考文献:

- [1] Hu Zhonggang et al., A sequel to the practical foundation course in vocal music, 43, 2015
- [2] 伊東 乾, 輝線から領域へ・・・声の周縁と非線形ポリフォニー, 「スペクトラル・コンダクティング」, 東京大学出版会, 2022 (in printing)
- [3] Hartmann, W. et al., Noise edge pitch and models of pitch perception, J.Acoust. Soc. Am. 145(4), 1993-2008, 2019

本研究は科学研究費「挑戦的研究(開拓)」20K20500「機械学習を用いた東アジア数理調和思想の実証的研究と共生倫理の検討 (研究代表者;筆者)」

<https://kaken.nii.ac.jp/ja/grant/KAKENHI-PROJECT-20K20500/> の助成を受けて行ったものです。関係ご各位に深く感謝します。