

4kHz以下の頭部伝達関数における後方知覚の手がかり*

◎中村風香（千葉工大・院），飯田一博（千葉工大・先進工）

1. はじめに

1960 年代以降，正中面音像定位の手掛かりについて多くの研究が進められ，頭部伝達関数(HRTF)の 5 kHz 以上の帯域の振幅スペクトルが重要であることが明らかにされている。

しかし，正中面音像定位のうち前後方向の判別においては，4 kHz 以下の成分だけでも可能であることが示されている[1]。また水平面においても，広帯域信号と比較すると若干精度は劣化するものの 4 kHz 低域通過ノイズでも音像定位が可能である[2]。したがって，4 kHz 以下にも何らかの前後方向知覚の手掛かりが存在すると考えられる。

本研究では，4 kHz 以下の帯域における後方知覚の手がかりを検討する。さらに，得られた手がかりの妥当性を音響心理実験により検証する。

2. 4 kHz 以下の前方と後方の HRTF の比較

4 kHz 以下の後方知覚の手がかりを探るため，成人男女 118 耳介の HRTF を用いて後方と前方のスペクトルレベルの差を求めた(Fig. 1)。各ドットは各耳介の(後方-前方)のレベルを表している。ほとんどの周波数で 0 dB 以下に布置されており，前方の音圧レベルが後方よりも大きい傾向がみられる。ただし，4 kHz 以上では個人差が大きい。一方，1 kHz 付近では後方の音圧レベルの方が大きく，個人差も比較的小さい(青の矢印)。

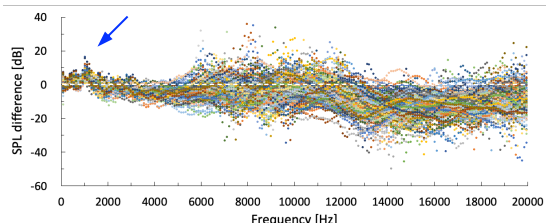


Fig. 1 Difference spectrum, subtracting sound pressure level (SPL) for HRTF for front direction from that for rear direction for 118 ears.

1 kHz 付近の後方と前方のレベル差を定量的に分析するため，118 耳介の後方と前方の HRTF のレベル差の平均値と 95 %信頼区間を

算出した(Fig. 2)。平均レベル差は 1031.25 Hz で最大となり，5.6 dB であった。95 % 信頼区間は 0.48 dB であった。

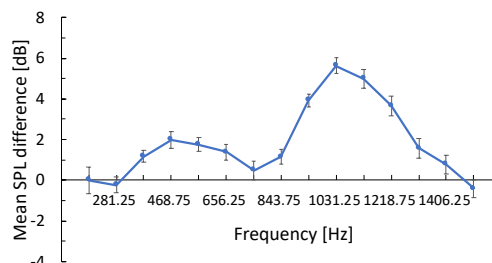


Fig. 2 Mean difference in SPL between rear and front HRTFs for 118 ears. Bars denote the 95% significance range.

Fig. 3 に 118 耳介において平均レベル差が最大となる周波数の相対頻度を示す。周波数分解能は 93.75 Hz である。全耳介のうち 49 % が 1031.25 Hz でレベル差が最大となり，86 % の耳介が 1031.25 Hz を中心とする 1/3 oct. の帯域でレベル差が最大となった。

以上より，後方の HRTF のレベルは前方と比較して 1 kHz 付近で卓越し，その個人差も小さいといえる。

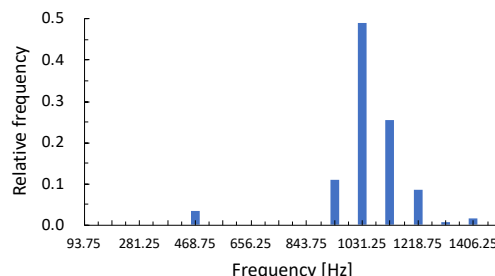


Fig. 3 Normalized histogram of frequencies for which SPL difference between rear and front HRTFs is maximum.

3. 1 kHz の卓越周波数帯域と方向決定帯域との関係

Blauert は正中面の前方，上方，後方に設置したスピーカから狭帯域信号(1/3 oct. バンドノイズ)を被験者に提示したところ，音源方向に関わらず特定の方向(前方，上方，後方)に音像を知覚する帯域が存在することを発見し，方向決定帯域と呼んだ[3]。

2 章で述べた 1 kHz の卓越周波数帯域(以降

* Cue for rear sound image localization in the head-related transfer function below 4 kHz, by NAKAMURA, Fuka and IIDA, Kazuhiro (Chiba Institute of Technology).

P0 と呼ぶ)は後方の方向決定帯域(800 Hz–1.6 kHz)と一致する。このことより、P0 は後方知覚の手がかりとして寄与することが期待されるが、一方で方向決定帯域は広帯域信号では上昇角知覚に影響を及ぼさないことも報告されている[4]。

そこで、P0 が後方知覚の手がかりとして寄与するか否かについて、次章で検証する。

4. P0 が後方知覚に及ぼす影響

P0 を除去した刺激を用いた音像定位実験により、P0 が 4 kHz 以下の帯域で後方知覚の手がかりとして寄与するか否かを検証した。

4.1. 実験方法

実験は防音室で実施した。刺激は以下の 3 種類を用いた。(1)被験者本人の後方の実測 HRIR と広帯域白色雑音(200 Hz–17 kHz)の畳み込み(WB-own), (2)被験者本人の後方の実測 HRIR と低域白色雑音(200 Hz–4 kHz)の畳み込み(LP-own), (3)P0 を除去した被験者本人の後方の実測 HRIR と低域白色雑音(200 Hz–4 kHz)の畳み込み(LP-P0-eliminated)。Fig. 4 に P0 を除去した本人 HRTF の 4 kHz 以下の振幅スペクトルの例を示す。

刺激の提示時間は 1.0 s で、3 種類の刺激を FEC ヘッドホン[5]とみなせるオープンエアヘッドフォン (SONY, MDR-MV1) により被験者に提示した。予備実験において、補正の有無で音像定位の精度に大きな違いがみられなかったため、ヘッドホンの伝達関数は補正していない。提示音圧レベルは 62 dB SPL である。

各刺激はランダムな順に 10 回ずつ提示した。被験者は正常な聴力を持つ成人 5 名(男性 2 名, 女性 3 名, A–E)で、マッピング法により音像の上昇角を回答した。被験者は各刺激を何回でも聴くことができるが、回答後はその刺激を聴くことはできない。

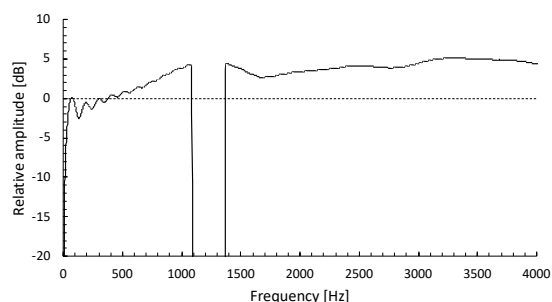


Fig. 4 Example of transfer function obtained by convolution of subject's HRIR at 180° and P0-eliminated filter.

4.2 実験結果

A 回答分布

各被験者の回答を Fig. 5 に示す。WB-own では、被験者 A, B, D は概ね目標方向の 180°に回答した。被験者 C は上昇角 130–165°の間に回答した。被験者 E は 0–90°に回答し、従来の研究[6]と一致しないため、以後の分析からは除外した。

LP-own の回答分布は従来の研究[1]と一致し、4kHz 以下の成分だけでも被験者 B, C, D は概ね目標方向の 180°に回答した。被験者 A は正中面の後ろ半分の 90–180°に回答した。

LP-P0-eliminated では、被験者 A, B, C, D の回答はいずれも上昇する傾向がみられた。

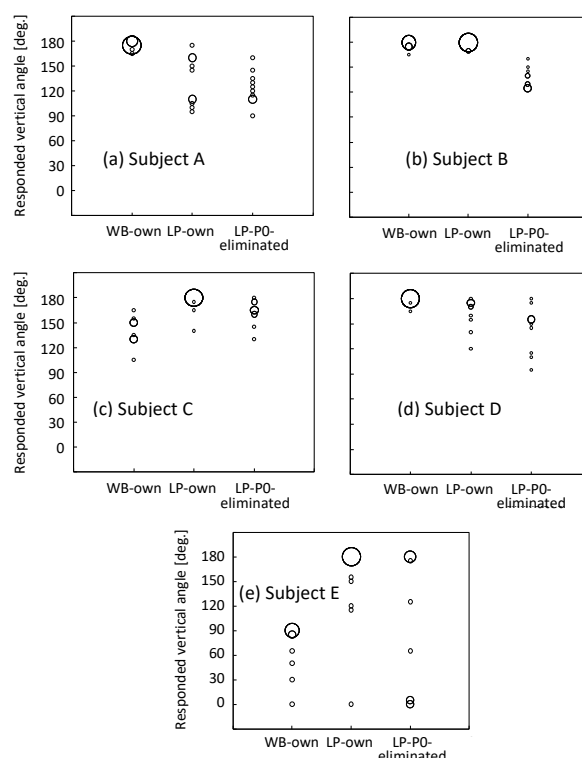


Fig. 5 Individual responses for subjects A–E to WB-own, LP-own, and LP-P0-eliminated stimuli at 180°.

B 平均仰角誤差

平均仰角誤差を Table 1 に示す。4 名の被験者の平均値は、WB-own で 12.4°, LP-own で 18.7°, LP-P0-eliminated で 38.1°であった。WB-own では被験者 A, B, D の誤差は 5°未満であった。いずれの被験者においても P0 を除去することにより仰角誤差は増加した。

Table 2 に LP-own と LP-P0-eliminated の平均仰角誤差に対する Mann-Whitney U 検定の結果を示す。4 名中 2 名で有意な差が認められた($p < 0.01$)。

Table 1 Mean elevation angle error [deg].

Subject	WB-own	LP-own	LP-P0-eliminated
A	4.5	48.5	56.1
B	3.5	2.0	42.4
C	39.3	6.2	18.2
D	2.5	18.2	35.7
Ave.	12.4	18.7	38.1

Table 2 Results of Mann-Whitney U test for LP-own and LP-P0-eliminated under 4 kHz.

**: $p < 0.01$, *: $p < 0.05$.

Subject			
A	B	C	D
	**	**	

C 前後誤判定率

被験者 A-D で前後誤判定は生じなかった。

D 頭内定位率

頭内定位率を Table 3 に示す。WB-own では被験者 A が 1 度頭内定位した。LP-own では被験者 A,D で頭内定位する割合が高く、4 名の平均値は 0.50 であった。被験者 B は LP-P0-eliminated の頭内定位率は LP-own より高く、被験者 A,C ではわずかに高い結果となった。

Table 3 Ratio of inside-of-head localization.

Subject	WB-own	LP-own	LP-P0-eliminated
A	0.10	0.90	1.00
B	0.00	0.10	0.60
C	0.00	0.00	0.10
D	0.00	1.00	1.00
Ave.	0.03	0.50	0.68

5. P0 がみかけの音源の幅に及ぼす影響

広帯域白色雑音とパラメトリック HRTF (N1N2P1P2 で構成)[7,8]を用いた従来の音像定位実験[9]の内観報告で、N1N2P1P2 に P0 を追加することにより (Fig. 6), みかけの音源の幅 (ASW: Auditory Source Width) が狭くなり、実測 HRTF の ASW に近づくことが指摘された。そこで、シェッフェの一対比較の浦の変法を用いて ASW を比較する実験を行った。

5.1 実験方法

実験は無響室で行った。HRTF は以下の 3 種類を用いた。(1)被験者本人の後方の実測 HRTF, (2)被験者本人の後方のパラメトリック HRTF (N1N2P1P2 で構成), (3)被験者本人の後方のパラメトリック HRTF (N1N2P1P2+P0 で構成)。音源信号は広帯域白色雑音 (200 Hz–17 kHz) で、提示時間は 1.0 s である。音源信号に HRIR を畳み込んだ刺激を FEC ヘッドホンとみなせる

オープンエアヘッドホン (beyerdynamic, DT990 PRO) により被験者に提示した。ヘッドホン伝達関数は補正していない。提示音圧レベルは 72 dB である。被験者は正常な聴力を持つ成人 5 名 (男性 2 名, 女性 3 名, A,B,F–H) である。

2 つの刺激を対にして提示し、先に提示された刺激に対して、後に提示された刺激の ASW が狭いか広いかを回答させた。非常に狭い場合は -2.0, どちらでもない場合は 0.0, 非常に広い場合は 2.0 とした。被験者は、-2.0 から 2.0 までの数直線に 0.1 間隔の目盛をつけた回答用紙にペンで印をつけて回答した。

刺激対は $3 \times 2 = 6$ 種類あり、各刺激対をランダムな順に 5 回ずつ被験者に提示した。

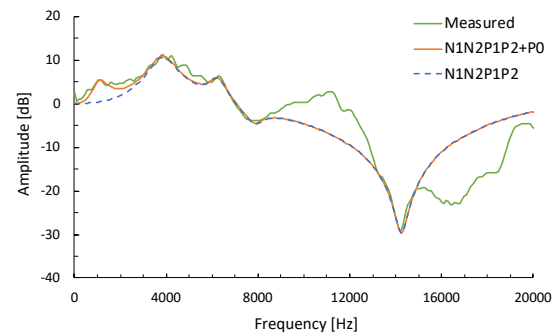


Fig. 6 Comparison among a measured HRTF and parametric HRTFs (N1N2P1P2 and N1N2P1P2 + P0) for the rear direction.

5.2 実験結果

各被験者の ASW の布置を Fig. 7 に示す。全被験者ともに ASW が最も広いのは N1N2P1P2 で、被験者 B を除く被験者は実測 HRTF が最も ASW が狭い。被験者 B では僅差ではあるが N1N2P1P2+P0 が最も ASW が狭い。

有意水準 5 % のヤードスティックと比較すると、被験者 H を除くすべての被験者で N1N2P1P2+P0 と N1N2P1P2 の ASW に有意な差が認められた。一方、N1N2P1P2+P0 と実測 HRTF の ASW には被験者 A,B,F では有意差があるとはみなせない。また、全被験者で N1N2P1P2 は実測 HRTF の ASW よりも有意に広い。

以上より、N1N2P1P2 の ASW は実測 HRTF よりも広がる (音像がぼやける) が、P0 を付加することによって実測 HRTF の ASW に近づくといえる。

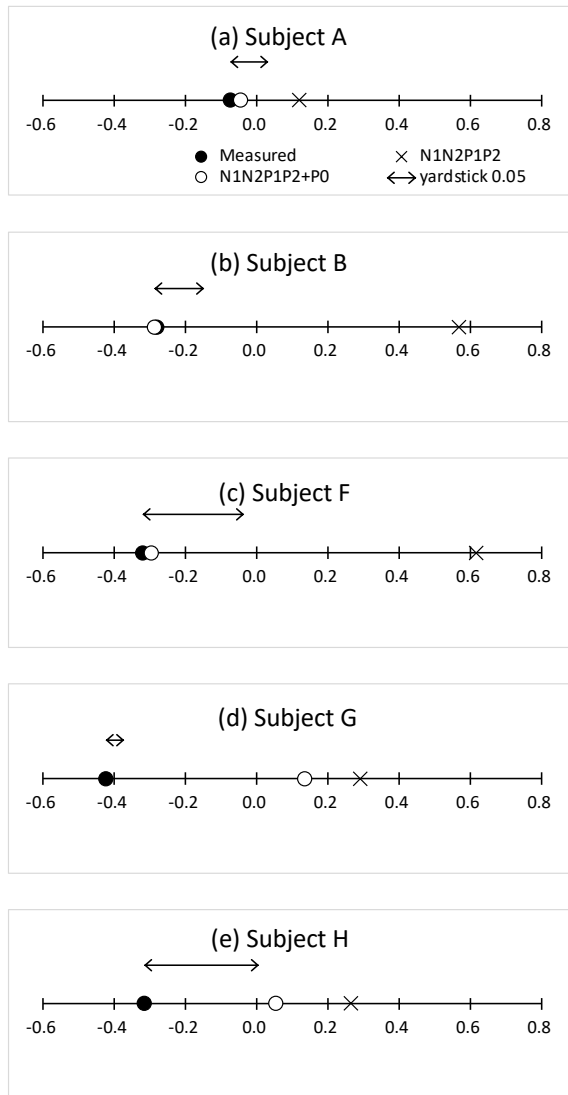


Fig. 7 Auditory source width for measured HRTF (●), N1N2P1P2 (×), and N1N2P1P2+P0 (○) for rear direction. Arrow indicates yardstick (0.05).

6. 考察

6.1 P0 の成因

P0 の成因について検討する。従来, FDTD 法で耳介形状のみを使用して算出した HRTF と, 耳介を含む頭部全体の形状を使用して算出した HRTF が比較されている[10]。前者では, 1 kHz 付近では上昇角に関わらず耳介の影響はみられなかった。波長と耳介寸法の関係から, これは妥当な結果と考えられる。一方, 後者では 1 kHz 付近で頭部は HRTF に影響を及ぼすが, 音源が前方の場合と後方の場合でその影響の違いはみられなかった。

したがって, 胴体が影響している可能性が高い。また, P0 は個人差が小さいことも, 胴体の影響を支持している。今後, 胴体を含んだ数値計算などで, さらに成因を検討する必要がある。

6.2 前方の音像定位の手がかり

4 kHz 以下の周波数での前方知覚の手がかりについて検討する。前方の手がかりの 1 つとして P1 が考えられる。P1 は正中面のいずれの方向においても実測 HRTF と同等の音像定位精度を提供する最小のパラメトリック HRTF の構成要素の中のピーク[6,7]である。また, P1 は前方の方向決定帯域と一致する。本研究の 118 耳介の P1 周波数の平均値は 4020.0 Hz で, 標準偏差は 249.5 Hz であった。平均値は約 4 kHz であるが, P1 の Q は 1-2 であり裾野の広いピークであることから, 4 kHz 以下の帯域にも前方知覚の手がかりとして影響を及ぼしていると考えられる。

7. おわりに

本研究では, HRTF の 4 kHz 以下の帯域における後方知覚の手がかりを検討した。その結果を以下にまとめる。

- (1) 118 耳介の前方と後方の HRTF の差を分析し, 1 kHz 付近で後方は前方よりも卓越することを示した。また, この卓越周波数帯域(P0)は Blauert の後方の方向決定帯域と一致した。
- (2) P0 が 4 kHz 以下の帯域で後方知覚の手がかりとして寄与するか否かを音像定位実験で検証した。P0 を除去することにより仰角誤差が増加し, P0 の寄与を支持する結果となった。
- (3) N1N2P1P2 で構成されたパラメトリック HRTF に P0 を付加することで ASW は狭くなり実測 HRTF の ASW に近づいた。

参考文献

- [1] 森本と斉藤, 日本音響学会聴覚研究会資料, H-40-1, 1977.
- [2] 森川と平原, 日本音響学会誌 68 巻 41, 71-179, 2012.
- [3] Blauert, ACUSTICA, 22, 205-213, 1969/70.
- [4] 竹内と飯田, 音講論(秋), 665-668, 2017.
- [5] Møller *et al.*, J Audio Eng. Soc., 43, 203-17, 1995.
- [6] K. Iida, In: Head-Related Transfer Function and Acoustic Virtual Reality. Springer; 2019. p.27.
- [7] K. Iida *et al.*, Appl. Acoust., 68, 835-850, 2007.
- [8] K. Iida and Y. Ishii, Appl. Acoust., 129, 239-247, 2018.
- [9] 中村と飯田, 日本音響学会聴覚研究会資料, Vol.53, No.6, H-2023-46, 2023.
- [10] H. Takemoto *et al.*, J Acoust. Soc. Am., 132, 3832-41, 2012.