

# 両耳間時間差・音圧差の相互作用による 音像の分離知覚の検知限

酒井翼<sup>1</sup> 森川大輔<sup>1</sup> モクタリ パーハム<sup>1</sup> 平原達也<sup>1</sup>

**概要**：ヒトは音色が同じ音であっても、音の到来方向の違いによって生じる両耳間時間差 (ITD) や両耳間音圧差 (ILD) を利用することで、音像を2つ以上に分離させて知覚することができる。しかし、その検知限と ITD と ILD が同時に作用する条件での音像の分離知覚がどのように生じるかはまだ明らかでない。そこで本稿では、ITD または ILD と、それらの相互作用によって生じる音像の分離知覚の検知限を明らかにするために、刺激音信号に ITD または ILD を与える条件と、頭部伝達関数を用いてバイノーラル信号を合成した条件の分離知覚実験を行った。その結果、ITD と ILD が同時に作用する場合の分離知覚には、ITD が与える影響が大きいと示唆された。

## Detection limit of sound images spatially split by interaction of interaural time and level difference

TSUBASA SAKAI<sup>†1</sup> DAISUKE MORIKAWA<sup>†1</sup> PARHAM MOKHTARI<sup>†1</sup>  
TATSUYA HIRAHARA<sup>†1</sup>

### 1. はじめに

ヒトは目的の音に注意を向けて聴くことでその音を聴き取りやすくなることが知られている。この能力はカクテルパーティー効果と呼ばれている[1, 2]。カクテルパーティー効果は音色の違い、言語の違い、注意の向け方、音の到来方向の違いなどの様々な手がかりが作用して生じていると知られている。しかし、これらの手がかりがどのように作用して生じているかはまだ明らかではない。

音の到来方向の違いによって生じる音響的な特徴は、両耳へ到達する音の時間が異なる両耳間時間差 (ITD : Interaural Time Difference)、両耳へ到達する音のレベルが異なる両耳間音圧差 (ILD : Interaural Level Difference) がある。ITD または ILD によって、目的音検知能力が上昇することについては多くの研究がなされている[3, 4, 5]。しかしこれらの報告は、目的音と雑音での音色の違いによる手がかりも影響している。音の到来方向の違いの手がかりがカクテルパーティー効果に与える影響を明らかにするためには、音色の違いによる手がかりが影響しない条件が好ましい。

森川は音色が同じ2つのホワイトノイズに ITD または ILD を与えると、音像が頭内で左右に2つに分離して知覚されることを明らかにした[6]。しかし、その検知限と、ITD と ILD が同時に作用する条件での音像の分離知覚については調べられていない。

そこで本稿では、ITD または ILD と、それらの相互作用によって生じる音像の分離知覚の検知限を明らかにするために、刺激音信号に ITD または ILD を付与した信号、および刺激音信号に頭部インパルス応答 (HRIR : Head-Related Impulse Response) を畳み込んで合成したバイノーラル信

号に対する音像の分離知覚実験を行った。

### 2. 実験方法

#### 2.1. 受聴者

受聴者は22歳から35歳の6名である。受聴者の最小可聴閾値は、125 Hz から 12 kHz で正常な聴力であることを確認した。

#### 2.2. 刺激音

刺激音信号は、帯域フィルタ (250 Hz ~ 8 kHz) を通した2つの無相関なピンクノイズ (PN1, PN2) である。刺激音信号に ITD または ILD を付与した信号と、受聴者本人の HRIR を畳み込んだ信号を刺激音として用いた。

実験条件は、刺激音信号に付与した ITD を変化させる条件 ITD-cond と付与した ILD を変化させる条件 ILD-cond、刺激音信号に畳み込む HRIR を正面から変化させる条件 F-cond と背面から変化させる条件 B-cond の4条件である。受聴者は ITD-cond と ILD-cond では、図 1a のように頭内で左右に分離した音像を知覚し、F-cond と B-cond では図 1b のよう頭外の正面や背面で左右に分離した音像を知覚する。バイノーラル信号に対して、F-cond と B-cond の2条件で実験を行ったのは、正面と背面でバイノーラル信号の ITD と ILD の組み合わせが異なるからである。

刺激音信号に付与した ITD を変化させる条件では、信号の配列を循環シフトして時間差を制御した。PN1 は Lch から Rch より ITD 分先の信号が、PN2 は Rch から Lch より ITD 分先の信号が出力される。循環シフトで ITD を与えているため、Lch と Rch で刺激音のオンセットは等しい。帯域フィルタを通したピンクノイズをヘッドホン (HDA-200, Sennheiser) で再生したときの音圧レベルを 70 dB とした。

<sup>†1</sup> 富山県立大学 工学部  
Faculty of Engineering, Toyama Prefectural University

刺激音信号に付与した ILD を変化させる条件では、信号の振幅を小さくし、音圧差を制御した。PN1 は Lch から Rch より ILD 分大きく、PN2 は Rch から Lch より ILD 分大きな音圧で出力される。音圧レベルは、刺激音信号に ILD が付与されていない信号で 70 dB とした。

刺激音信号に畳み込まれる HRIR を変化させる条件では、受聴者個人の HRIR を信号に畳み込み、バイノーラル信号を合成した。HRIR は水平面 1.5 m の円周上、5 度間隔で測定されたものを 1 度間隔に補間[7]して用いた。用いた HRIR の仰角はすべて 0 度である。HRIR はサンプリング周波数 48 kHz で測定されたが、時間分解能を上げるために、192 kHz にアップサンプリングした。音圧レベルは、水平角 0 度の HRIR を畳み込んだ信号で 70 dB とした。

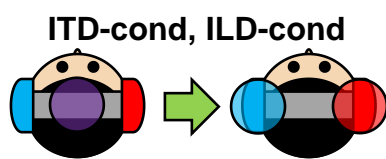


図 1a 受聴者が知覚する音像の模式図 (ITD or ILD-cond)

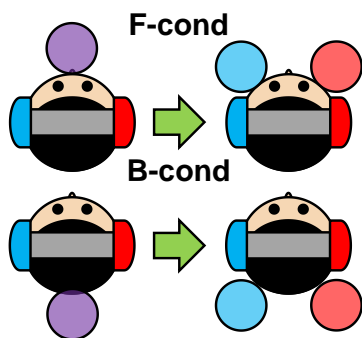


図 1b 受聴者が知覚する音像の模式図 (F-cond, B-cond)

### 2.3. 実験条件

実験は図 2 に示す実験系で行った。PC (Windows10, 64bit) で生成した刺激音信号を DA 変換機 (Fireface UCX, RME), ヘッドホンアンプ (AT-HA21, audio-technica) を通し、ヘッドホン (HDA-200, Sennheiser) で再生した。サンプリング周波数は 192 kHz とした。

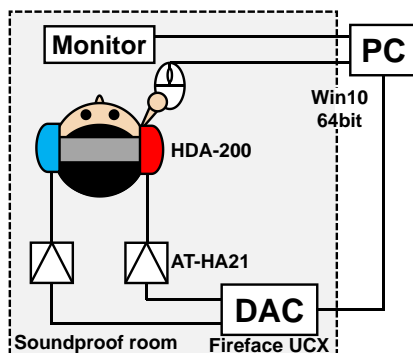


図 2 実験系

### 2.4. 実験手続き

実験は調整法を用いて行った。受聴者はマウスホイールを用いて刺激音を制御し、音像が図 1a, b に示すように分離したところを回答した。各条件で 10 回の回答を得た。そのうち 1~5 回目の回答を練習試行、6~10 回目の回答を本実験とした。

受聴者はマウスホイールを用いて刺激音の ITD, ILD または畳み込まれる HRIR を連続的に制御した。ITD-cond の ITD の初期値は  $0 \mu\text{s}$ , ILD-cond の ILD の初期値は 0 dB, F-cond の水平角の初期値は 0 度, B-cond の水平角の初期値は 180 度である。ホイール 1 スクロールあたり ITD は  $\pm 1$  サンプル (約  $5 \mu\text{s}$ ), ILD は  $\pm 0.2 \text{ dB}$ , HRIR は  $\pm 1$  度切り替わる。

### 3. 実験結果

図 3 に ITD-cond と ILD-cond の受聴者ごとの音像の分離知覚の検知限 ITD-DL, ILD-DL に示す。受聴者 1, 2, 3, 6 は先行研究[8]の 9 名のうちの 4 名である。受聴者 4 は、ILD-cond で分離知覚が生じなかった。

図 4 に F-cond と B-cond の受聴者ごとの音像の分離知覚の検知限 F-DL<sub>θ</sub> と B-DL<sub>θ</sub> を示す[9]。DL<sub>θ</sub> は 2 つの刺激音信号に畳み込まれている HRIR の角度差である。受聴者 2 と 5 のあるセッションの回答を確認すると、受聴者 2 は HRIR の水平角を上下法のように調整し、回答していた。この傾向は受聴者 1, 3, 4, 6 も同様である。一方、受聴者 5 は HRIR の水平角を極限法のように調整しており、検知限より大きい水平角の HRIR が畳み込まれた刺激音を受聴していかった。また、受聴者 5 以外の受聴者の回答時間は約 13 ~ 26 秒で、受聴者 5 の回答時間は約 4 秒であった。

ILD-cond で分離が生じなかった受聴者 4 と、F-cond と B-cond の F-DL<sub>θ</sub>, B-DL<sub>θ</sub> が極めて小さく、回答時間も短い受聴者 5 を除外した受聴者 1, 2, 3, 6 の 4 名で各結果の平均値を算出すると、ITD-DL は  $384 \mu\text{s}$ , ILD-DL は 17 dB, F-DL<sub>θ</sub> は 41 度, B-DL<sub>θ</sub> は 44 度であった。

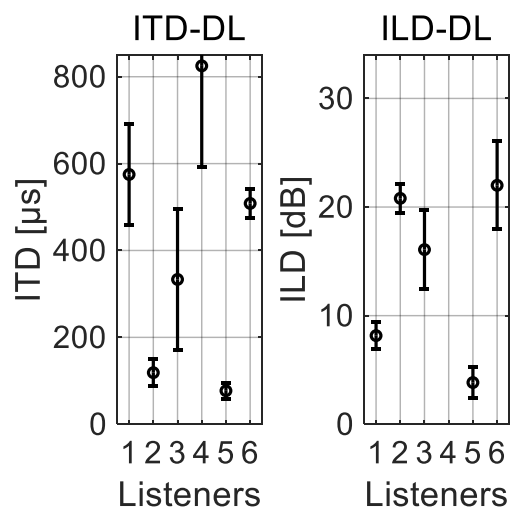


図 3 被験者ごとの ITD-DL と ILD-DL

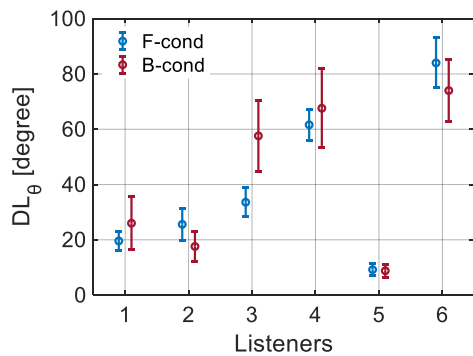


図4 被験者ごとの F-DL<sub>θ</sub> と B-DL<sub>θ</sub>

#### 4. 考察

音像が分離したときの HRIR を畳み込んだ 2 つの刺激音のそれぞれの ITD の差 ( $DL_{\theta}^{ITD}$ ), ILD の差 ( $DL_{\theta}^{ILD}$ ) を算出した. 図 5 に受聴者 1, 2, 3, 6 の ITD-DL, ILD-DL, F-DL<sub>θ</sub><sup>ITD</sup>, F-DL<sub>θ</sub><sup>ILD</sup>, B-DL<sub>θ</sub><sup>ITD</sup>, B-DL<sub>θ</sub><sup>ILD</sup> の平均値を示す.

ITD-DL の平均値は 384 μs, F-DL<sub>θ</sub><sup>ITD</sup> の平均値は 384 μs, B-DL<sub>θ</sub><sup>ITD</sup> の平均値は 407 μs であった. 条件 (ITD-cond, F-cond, B-cond) について 1 要因の分散分析を行った結果, 条件の要因に有意差はなかった [ $F(2, 57) = 0.07, p > 0.5$ ].

ILD-DL の平均値は 17 dB, F-DL<sub>θ</sub><sup>ILD</sup> の平均値は 11 dB, B-DL<sub>θ</sub><sup>ILD</sup> の平均値 7 dB であった. 条件 (ILD-cond, F-cond, B-cond) について 1 要因の分散分析を行った結果, 条件の要因に有意差があった [ $F(2, 57) = 17.2, p < 0.001$ ].

ITD-DL, F-DL<sub>θ</sub><sup>ITD</sup>, B-DL<sub>θ</sub><sup>ITD</sup> は, 条件間の要因に有意差はなかった. 一方 ILD-DL, F-DL<sub>θ</sub><sup>ILD</sup>, B-DL<sub>θ</sub><sup>ILD</sup> は, 条件間の要因に有意差があった. したがって, ITD と ILD が同時に作用する場合の音像の分離知覚には, ITD が与える影響が大きいと考えられる.

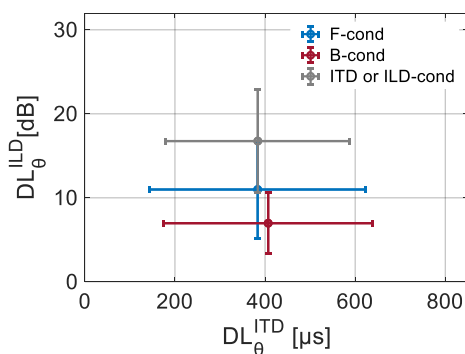


図5 各条件の DL<sub>θ</sub><sup>ITD</sup> と DL<sub>θ</sub><sup>ILD</sup>

#### 5. まとめ

本稿では, ITD または ILD と, それらの相互作用によって生じる音像の分離知覚実験を行った. その結果, ITD-DL が 384 μs, ILD-DL が 17 dB, F-DL<sub>θ</sub> が 41 度, B-DL<sub>θ</sub> が 44 度, F-DL<sub>θ</sub><sup>ITD</sup> が 384 μs, F-DL<sub>θ</sub><sup>ILD</sup> が 11 dB, B-DL<sub>θ</sub><sup>ITD</sup> が 407 μs, B-DL<sub>θ</sub><sup>ILD</sup> が 7 dB であった. ITD-DL, F-DL<sub>θ</sub><sup>ITD</sup>, B-DL<sub>θ</sub><sup>ITD</sup> は一致し, ILD-DL, F-DL<sub>θ</sub><sup>ILD</sup>, B-DL<sub>θ</sub><sup>ILD</sup> は異なることから, 合成バイノーラル信号の音像の分離知覚には主に ITD が寄与していると考えられる.

#### 謝辞

本研究の一部は科研費 (20K19828) および東北大学電気通信研究所における共同プロジェクト研究の支援を受けた.

#### 参考文献

- [1] E. C. Cherry, "Some experiments on the recognition of speech with one and two ears," *J. Acoust. Soc. Am.*, 25, 975-979 (1953).
- [2] B. G. Shinn-Cunningham, "Object-based auditory and visual attention", *Trends Cogn. Sci.*, 12, 182-186 (2008).
- [3] M. Ebata, T. Sone, and T. Nimura, "Improvement of hearing ability by directional information," *J. Acoust. Soc. Am.*, 43, 289-297 (1968).
- [4] K. Saberi, L. Dostal, T. Sabralodabei, V. Bull, and R. D. Perrot, "Free-field release from masking," *J. Acoust. Soc. Am.*, 90, 1355-1370 (1991).
- [5] K. Kuroda, J. Li, Y. Iwaya, M. Unoki, and M. Akagi, "Effects of spatial cues on detectability of alarm signals in noisy environments," *PRINCIPLES AND APPLICATIONS OF SPATIAL HEARING* (World Scientific, Singapore, 2011), 484-493.
- [6] 森川, "両耳間差による音像の分離知覚," 信学技報 114(242), SP2014-83, 43-48, 2014.
- [7] 西野ら, "水平面上の頭部伝達関数の補間," 日本音響学会誌 55(2), 91-99, 1999.
- [8] 酒井ら, "両耳間時間差・音圧差による音像の分離知覚の検知限," 音講論(秋), 577-578, 2020.
- [9] 酒井ら, "合成バイノーラル信号の音像の分離知覚の検知限," 音講論(春), 641-642, 2021.