

# 両耳間時間差の脳内処理メカニズムへの心理物理学のアプローチ Psychophysical approach to understanding interaural-time-difference processing in the brain

古川 茂人 (日本電信電話株式会社 NTT コミュニケーション科学基礎研究所)  
Shigeto FURUKAWA (NTT Communication Science Labs, NTT corporation)

furukawa.shigeto@lab.ntt.co.jp

**Abstract**—This paper reviews three types of psychophysical experiments conducted in the author's laboratory, aiming at understanding interaural-time-difference (ITD) processing in the human auditory system. The first experiment, on the basis of the signal detection theory, indicated that the process for ITD is different between high and low frequency regions: In the low frequency regions, there are partially independent processes for ITD and ILD (interaural level difference), whereas in the high frequency regions, a common mechanism likely processes the two cues. The second experiment was concerned with the hierarchy of ITD information processing and the temporal limit that a mechanism at each level of processing can track changes in ITD. The results indicated that the ITD processor can track changes in ITD up to a rate of 20 Hz and that the loss of tracking ability of the auditory system for a rate slower than that is attributable to a higher level mechanism, which might include a process which integrates ITD and ILD information. In the third experiment, listeners were trained in a pitch discrimination task. The task required the listeners effectively use the information of temporal fine structure of the stimuli, the information being critically important also in ITD processing. The pitch training deteriorated the listeners' performance in the (untrained) ITD discrimination task. The result implies that the pitch and ITD processes compete with each other for limited neural resources.

## 1. はじめに

耳に届く音の両耳間時間差(interaural time difference, ITD)は両耳間レベル差(interaural level difference, ILD)とならんで、音源定位の主要な手がかりである。競合音が存在する実環境で、目的とする音を聞き取る際にも、この ITD は重要な役割を果たすと考えられている。これまでにメンフクロウ、ネコ、スナネズミなどの動物モデルを用いた生理学実験により、聴覚系における ITD 処理機構の概要がある程度明らかになってきている (例えば[1-3])。

しかしながら、ヒトはこれらの動物モデルとは、身体の形状、可聴周波数範囲、生態学的な位置づけが異なっている。神経系が身体的、生態学的な条件に合わせて進化・発達しうることを考えると、これら動物モデルから得られた結果をそのままヒトに適用できるとは限らない。ヒトを対象にした実験では、動物に対してのように侵襲的な生理実験を行うことができない。また、既存の脳機能イメージング手法は、空間・時間分解能が十分でなく、ITD 処理が行われる脳幹内の微小な神経核の活動を明らかにする

ことはできない。

このような状況では、心理物理学のアプローチは、ヒトの ITD 処理メカニズムを探るうえで有効である。本稿では、ITD の脳内処理メカニズムの解明のため、特に脳内処理モジュールの構成や独立性に着目して筆者の研究室で実施された 3 種類の心理物理実験を紹介する。

## 2. ITD と ILD の処理メカニズムの独立性・周波数間の違い[4]

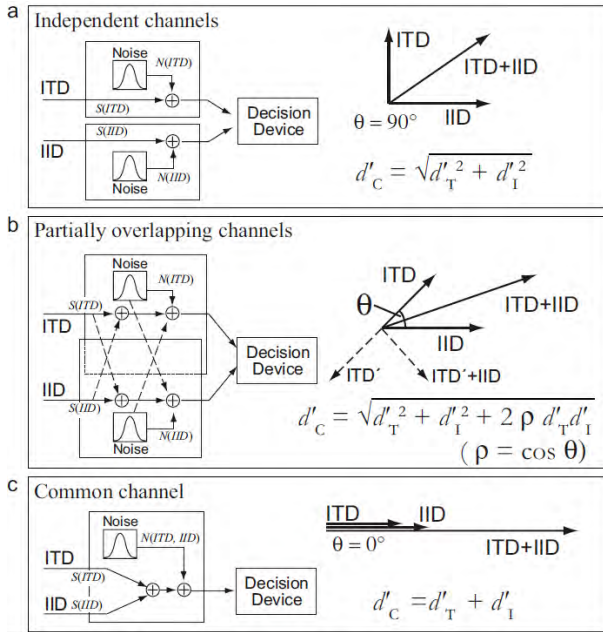
動物モデルを用いた生理実験により、脳幹にはそれぞれ ITD と ILD の処理に特化した神経核 (それぞれ、上オリーブ内側核および外側核; medial superior olive (MSO) および lateral superior olive (LSO)) が存在することが示されている[2]。これは、ITD と ILD が独立した機構で処理されている可能性を意味するものである。

しかし、実際にヒトの聴覚系において、ITD と ILD の処理がどの程度独立かは明らかではない。さらに、ITD および ILD の情報が広い周波数帯域にわたって同じメカニズムで処理されているとは限らない。本研究では、ITD および ILD 処理に関わる仮想的な「チャンネル」を仮定し、信号検出理論[5]の枠組みに従ってチャンネルの重なり・独立性を評価し、周波数帯域間で比較した。

信号検出理論[5]では、各チャンネルの出力は「信号」(手がかりの強度に関連した活動量)とそれに加わる「内部ノイズ」の 2 つの要素によって特徴付けられる。心理物理実験によって計測される、刺激音の検出感度(正答率等から導出される  $d'$  値)は、この信号対ノイズ比によって決定されると考える。ITD または ILD がそれぞれ単独で変化した場合には、それぞれに対する検出感度  $d'_T$  および  $d'_L$  は、対応するチャンネルでの信号対ノイズ比で表されると考える。ITD と ILD が同時に(同方向に)変化するときの感度  $d'_C$  は、チャンネルの独立性に依存する。各チャンネルからの出力が、線形かつ最適に組み合わせられると仮定すると、 $d'_C$  は、

$$d'_C = \sqrt{d'_T{}^2 + d'_L{}^2 + 2\rho d'_T d'_L}$$

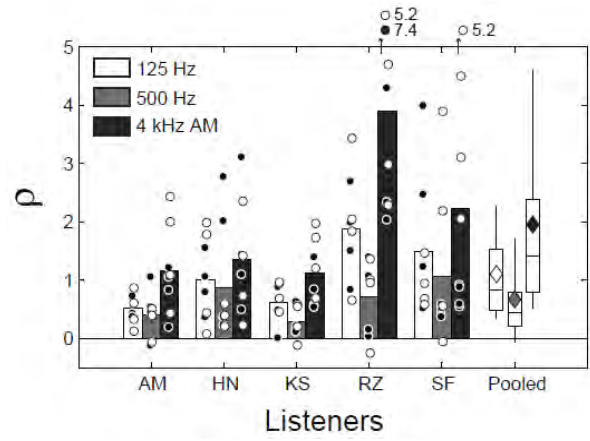
で予測される (Fig. 1) [6, 7]。ここで  $\rho$  はチャンネルの重なり程度を表す指標 (内部ノイズのチャンネル間相関係数) で、値が大きいほどチャンネルの重なりが大きく、0 (チャンネル同士が完全に独立) から 1



**Fig. 1.** Illustration of the working model used for evaluating cue interaction [4]. The model assumes channels for ITD and ILD processing, that overlap each other to various degrees. Each channel has internal noise, which is added to the signal that is related to cue strength. The magnitudes of the internal noise and the signal in the ITD and ILD channels depend on the input ITD and ILD, respectively. The outputs from the channels are added linearly. The decision device was assumed to base its judgments on changes in the weighted linear sum. It was also assumed that the weights on the channel outputs are chosen to optimize the performance. a. An extreme case in which the two channels are completely independent. Each of the channels has independent noise sources. In this case, the dimensions for ITD and ILD are thought of as orthogonal, and the  $d'$  value for simultaneous and consonant changes of ITD and ILD would be the root mean square sum of the  $d'$  values for individual changes. b. The two channels partially overlap. Fractions of the signal and the noise in one channel are added mutually to the other channel. The degree of channel overlap can be expressed as the angle of the ITD and ILD dimensions,  $\theta$  (see the vector diagram of solid-line arrows). The  $\cos \theta$  value is equal to the correlation coefficient  $\rho$  for the noise at the outputs of the two channels. c. ITD and ILD information is combined linearly in a single channel, in which the common internal noise is added to the signal. Equivalently, ITD and ILD are represented by a single dimension. The  $d'$  value for combined ITD and IID changes is the simple sum of the  $d'$  values for individual changes.

(ITD と ILD が共通の単一のチャンネル内で処理される)の間の値をとる。 $d'_T$ , および  $d'_L$  が一定ならば、 $\rho$  が大きくなるほど、 $d'_C$  も大きくなる。逆に、心理物理実験から得られた  $d'$  値を上式に当てはめ、 $\rho$  を推定することにより、チャンネル間の独立性を推測できる。

心理物理実験では、両耳手がかりの変化 (ITD のみ、ILD のみ、ITD・ILD 同時) に対する検出感度 (それぞれ  $d'_T$ ,  $d'_L$ ,  $d'_C$ ) を測定した。刺激音として次の 3 種を用いた。低周波数刺激として、周波数が 125 Hz および 500 Hz の正弦波音 (125-Hz tone, 500-Hz tone)、および高周波数刺激として、搬送周波数が 4 kHz の振幅変調音 (4-kHz AM tone) である。4-kHz AM tone の変調波は、125 Hz 正弦波を半波整流し、低域



**Fig. 2** The  $\rho$  values estimated for the three stimulus types. For each stimulus type, individual  $\rho$  estimates for 4  $\Delta$ ITD/ $\Delta$ IID combinations and 2 ITD/IID change directions are shown by the symbols. A bar indicates the mean of the eight  $\rho$  estimates. The 10, 25, 50, 75, and 95th percentiles of the  $\rho$  estimates pooled for all listeners and conditions are shown by the box-and-whisker plots.

濾過 (<2000 Hz) したものである。この刺激は、4 kHz に同調した聴神経を刺激しているながら、神経の発火パターンが、125 Hz の純音に対する 125 Hz に同調した聴神経の発火パターンと実質的に同様になるように設計されている [8]。これらの刺激を、所望の ITD または ILD となるように、ヘッドホンで両耳に呈示した。

検出感度の測定には 2 区間 2 肢強制選択法を用いた。各観察区間は連続した 3 つのトーン・バーストからなつた。Reference 区間では、3 つのバーストの ITD, ILD は一定とした (Baseline ITD, ILD = 0  $\mu$ s, 0 dB)。Signal 区間では、2 番目のバーストについて、Baseline 値より  $\Delta$ ITD または  $\Delta$ ILD だけ外れた ITD または ILD とした。Reference および Signal 区間の順序は、試行ごとにランダムに変え、被験者には、2 つの観察区間を比較して、Signal 区間を示すように指示した。ITD と ILD を同時に変化させた場合、その方向は常に同一であった。つまり、右耳刺激が左耳に対して時間的に先行する ITD 変化には、右耳刺激レベルが相対的に増加する ILD 変化が伴つた。

数種類の  $\Delta$ ITD と  $\Delta$ ILD の組み合わせについて、それぞれ 210 試行の結果を得た。その結果から、検出感度を  $d' = \sqrt{2} \cdot z\text{-score}(P_c)$  によって算出した ( $P_c$  は正答率) [9]。こうして得られた  $d'$  値を上式に当てはめ、最小二乗法により各条件における  $\rho$  を推定した。

推定された  $\rho$  値を Fig. 2 に示す。それぞれの点は、 $\Delta$ ITD と  $\Delta$ ILD の各組合せに対して求められた  $\rho$  値で、それらの平均値が棒グラフで示されている。被験者間での推定値のばらつきは大きい、刺激種の系統的な効果も顕著である。2 要因分散分析 (要因 1: 刺激種; 要因 2: 被験者と変化方向の組合せ) の結果、刺激種の主効果は有意であった ( $p < 0.001$ ,  $F(2,$

90)=20.08)。刺激種間の多重比較では、4-kHz AM tone における  $\rho$  値が、他の刺激種よりも有意に大きいことが分かった(Tukey-Kramer test,  $p < 0.01$ )。低周波刺激同士の差は有意ではなかった( $p > 0.05$ )。これらの結果は、ITD および ILD の処理過程が、周波数間で異なることを示唆するものである。

4-kHz AM tone における  $\rho$  値の平均は、1.96 であった。これは、本研究で想定した  $\rho$  値の範囲(最大値 1)を超える。これは、設定した仮定のいずれかが成立しないことを意味している。実際、観測された  $d'$  値は、 $\rho = 1$  (単一チャンネル) を仮定した場合の予測値よりも有意に大きかった(paired  $t$ -test:  $p < 0.001$ )。その理由として、ITD と ILD の非線形な加算が可能性として考えられるが、その実態は不明である。

125-Hz tone においては、 $\rho$  の平均値は 1.11 であり、共通チャンネル( $\rho = 1$ )による予測と統計的な差異はなかった。500-Hz tone については、 $\rho$  の平均値(0.67)は 0 と 1 の間の値をとった。これは、ITD および ILD チャンネルが部分的に重なる (ITD と ILD を処理する独立した過程がある) ことを示唆するものである。

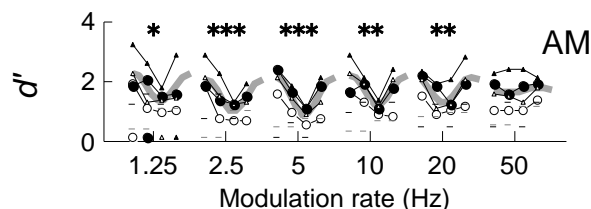
### 3. ITD の時間的変化に対する追従特性[10]

両耳機構は、緩慢(sluggish)であるといわれている。つまり、ITD または ILD が変動した場合、比較的ゆっくりとした (< 10Hz) ものであっても、その変化(変調)に追従できないとされている[11]。本研究では、その緩慢性の原因が、聴覚系のどのプロセスにあるかを検討した。

この検討にあたって、両耳機構を単純な 3 段階のプロセスとしてモデル化した。第 1 段階は、ITD および ILD の 2 つの手がかり (cue) を別々に処理する独立の機構から構成される (cue の個別処理レベル)。これら 2 つの機構は、それぞれ脳幹の上オリーブ内側核および外側核に相当すると考えても良い。第 2 段階では、第 1 段階の 2 つの機構からの出力を統合する (cue の統合レベル)。ここで、ITD および ILD の 2 つの cue が、加算され、あるいは互いに打ち消しあう。第 3 段階は、それ以上の高次機構をまとめたもので、判断機構(decision device)もそれに含まれる (高次レベル)。

あるプロセスにおいて「変調に追従できない」状態とは、「変調の方向に関する情報が失われている」状態と考えることができる。本研究では、変調方向に関する情報が、cue の個別処理のレベル(第 1 段階)、または cue の統合レベル(第 2 段階)以降の、いずれの段階で失われるかを検討した。実験では、刺激音の ITD および ILD を同時に正弦波状に変調して、その変調の検出感度を計測した。そして、ITD と ILD 変調の相対位相の効果を調べた。

もし、検出成績が、相対位相に依存するならば、



**Fig. 3.** The  $d'$  for detecting ITD and/or ILD modulations. The data are sorted according to the modulation rate (horizontal axis) and the relative phase of the ITD and ILD modulations, when the ITD and ILD were modulated simultaneously (connected symbols; relative phases of 0, 90, 180, and 270 degrees from left to right at each rate). Cases with statistically significant effects of relative phase are marked with asterisks above the plots (one-way repeated measures ANOVA; \*:  $p < 0.05$ ; \*\*:  $p < 0.01$ ; and \*\*\*:  $p < 0.001$ ). The gray curve with each listener and rate is a cycle of sinusoid fitted to the mean data across the combinations of  $\Delta$ ITD and  $\Delta$ ILD.

それは、変調方向の情報が、cue の個別処理のレベルで保持されている証拠ととらえることができる。これは、次のように考えると理解しやすい。ITD と ILD が同時に変調される時、cue の個別処理レベルの出力としての ITD と ILD の変調が同位相 (たとえば、左耳と比較して、右耳への刺激が時間的に先行するタイミングと音圧レベルが増加するタイミングが同期する) ならば、統合レベルにおいて 2 つの cue は加算され、逆位相ならば同レベルにおいて打ち消しあうであろう。このため、同位相の変調は検出しやすく、逆位相は検出しにくいことが予測される。もし、変調方向の情報が早い段階で失われるならば、このような、位相関係に応じた加算や打ち消しは起こらない。

実験では、中心周波数が 500 Hz の 1/3 オクターブバンド雑音を刺激音として用い、ヘッドホンを通して両耳に呈示した。被験者の検出感度は 2 区間 2 肢強制選択法を用いて計測した。2 つの刺激呈示区間のうち、「信号あり」区間では、刺激音の ITD および ILD を個別または同時に、0  $\mu$ s および 0 dB を中心とした正弦波状に変調した。変調周波数として 1.25, 2.5, 5, 10, 20, および 50 Hz を用いた。各変調周波数について、ITD 変調の、ILD 変調に対する位相遅延として、0, 90, 180, および 270 度を試した。「信号なし」区間では、変調は行わなかった (つまり diotic な刺激を両耳に呈示した)。各変調周波数について、ITD および ILD 変調の振幅 ( $\Delta$ ITD,  $\Delta$ ILD) に関して数首里の組み合わせ 20 種類の条件を試した: ITD と ILD の同時変調について 16 種類 (2 種類の  $\Delta$ ITD x 2 種類の  $\Delta$ ILD x 4 種類の位相遅延) および、ITD または ILD 単独の変調について 4 種類 (2 種類の  $\Delta$ ITD + 2 種類の  $\Delta$ ILD)。それぞれの条件において、81~117 試行を行った (被験者および変調周波数ごとに異なる)。得られた正答率より、 $d'$  値を算出した。

例として、1 名の聴取者から得られた結果を Fig. 1

に示す。各変調周波数（横軸）について、ITD と ILD の同時変調に対する検出感度  $d'$  値を示している。各変調周波数について、折線で結ばれた点は各位相遅延に対する結果である（左から、0, 90, 180, および 270 度）。20 Hz までの変調周波数に対して、位相遅延の有意な影響がみられた。この聴取者では、検出感度が最大となるのは、位相遅延が  $0^\circ$ （同位相）の時で、 $180^\circ$ （逆位相）で最小となる傾向が見られた。検出感度が最大となる位相遅延は、聴取者によって異なった。しかし、いずれの場合であっても、位相遅延の影響が見られたのは 20 Hz までの変調周波数であった。

前述のモデルから考えると、この結果は、ITD を処理するモジュールは、少なくとも 20 Hz までの変調に追従することができることを意味する。聴覚系全体で考えたときの追従の上限がその周波数よりも低いならば、その要因は、ITD と ILD を統合するモジュール以上のレベルの処理にあるといえる。

#### 4. ピッチの知覚学習の ITD 感度に与える影響[12]

ITD の処理には少なくとも 2 つのメカニズムが関与している。1 つは音の詳細な時間波形 (temporal fine structure, TFS) を、聴神経の位相固定 (phase locking) によって神経の時間的発火パターンへと変換し、中枢へと伝達するメカニズム (TFS 伝達メカニズム) である。もう 1 つは、前記のメカニズムによって伝えられる神経発火のタイミングを、両耳間で比較するメカニズム (両耳メカニズム) である。ITD の感度は、これら 2 つのメカニズムの精度によってある程度決定されると考えられる。このうち、TFS 伝達メカニズムは、ITD 以外の情報処理にも重要な役割を果たす。例えば、神経発火の時間間隔に基づいて TFS の周期性を計算することによって、ピッチを知覚することができると考えられる。

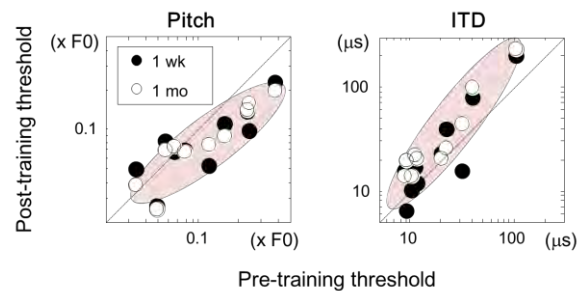
一般に、ある特定の知覚タスクを繰り返し行くと、その知覚属性について特異的に感度が向上する（知覚学習が起こる）ことがある。これは、このタスクに関わる神経処理モジュールが可塑的であることを意味する。この場合、この神経処理モジュールの出力を利用する他のタスクの成績も向上するであろう。逆に言えば、一見異なる知覚属性に関するタスクであっても、一方に関する知覚学習の結果が、もう一方のタスクの成績に影響するのであれば、これら 2 つの知覚属性は共通のモジュールを利用していると推察することができる。このパラダイムに則り、これまで多くの有効な知見が得られている（例えば [13]）。本研究では、TFS に基づくピッチタスクについての知覚学習が、ITD タスクに影響を与えるかを調べた。

刺激音は、帯域制限された調波複合音（基本周波

数 100 Hz ; 7 から 14 倍音成分を含む）である（H 音とする）。この音のすべての成分を、同時に一定の周波数 ( $\Delta f$ ) だけ高周波側にシフトさせると、振幅包絡（周期 100 Hz）を一定に保ったまま、その詳細な時間波形 (TFS) のみが変わる音（S 音とする）を作ることができる。この音は、H 音よりピッチが高い音として知覚される。聴覚末梢においては、その周波数分解能の制約のため、各周波数成分は分解されない。このため、H 音と S 音のピッチの違いは、TFS の違いによるものといえる[14]。実験では、2 種類の系列 (HHHH および HSHS) の音をランダムな順序で提示し、2 区間 2 肢強制選択法において 2 つの系列を弁別するタスクを聴取者に課した。このタスクでは、適応法によって  $\Delta f$  を変化させ、弁別閾値を求めた。聴取者としては、過去に聴覚心理学実験を経験したことのないものを採用した。これを 2 群に分け、一方の群（訓練群）では、このタスクを 12 日間行う訓練を課した。もう一方の群（コントロール群）では、同等な期間に訓練を行わなかった。訓練期間の前後には、同じピッチタスクに加えて、H 音を刺激として用いた ITD タスク、ILD タスク、および強度弁別タスクを行った。これらのタスクで、それぞれ ITD 検出、ILD 検出、および強度弁別閾値を計測した。

Fig. 4 は、訓練群について、訓練前後のピッチ弁別閾値を比較したものである。期待されたとおり、訓練後の閾値は、訓練前のそれから有意に減少した（つまり、成績が向上した）。しかし、ITD タスクについて、ピッチタスク訓練前後の閾値を比較すると、訓練後の閾値は有意に上昇した（成績が悪化した）。他のタスクについては、訓練前後の閾値変化は認められなかった。また、コントロール群については、訓練期間（実際は訓練はしていないが）前後の閾値の変化はいずれのタスクでも認められなかった。

ピッチの知覚学習による ITD 感度の悪化は、期待されていなかった結果である。これは、ピッチ情報と ITD 情報が、一つの神経モジュール内で同時に処



**Fig. 4.** Comparisons of post-training and pre-training thresholds. Filled and open symbols indicate thresholds obtained 1 week and 1 month after the pitch training, respectively. Each symbol represents one listener.

理されると仮定することで説明できるかもしれない。その神経モジュールの計算リソースは限られており、ピッチ処理と ITD 処理は競合する。ピッチタスクの訓練によって、より多くのリソースがピッチ処理に割かれることにより、その結果、ITD 処理のためのリソースが減少したのかもしれない。

## 5. おわりに

本稿をとおして、両耳情報の神経処理機構を推定するためのツールとして、心理物理学が有効であることを感じていただければ幸いである。

## 謝辞

知覚学習に関する研究は、鷺沢史歩氏と柏野牧夫氏との共同研究である。

## 参考文献

- [1] M. Konishi, "Coding of auditory space," *Annu Rev Neurosci*, vol. 26, pp. 31-55, 2003.
- [2] T. C. Yin, "Neural mechanisms of encoding binaural localization cues in the auditory brainstem," in *Integrative functions in the mammalian auditory pathway*, D. Oertel, *et al.*, Eds., ed New York: Springer, 2002, pp. 99-159.
- [3] B. Grothe, "Sensory systems: New roles for synaptic inhibition in sound localization," *Nat Rev Neurosci*, vol. 4, pp. 540-50, Jul 2003.
- [4] S. Furukawa, "Detection of combined changes in interaural time and intensity differences: Segregated mechanisms in cue type and in operating frequency range?," *J Acoust Soc Am*, vol. 123, pp. 1602-17, Mar 2008.
- [5] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. New York: Krieger, 1974.
- [6] D. M. Green, "Detection of multiple component signals in noise," *J. Acoust. Soc. Am.*, vol. 30, pp. 904-911, 1958.
- [7] E. R. Hafter, *et al.*, "The combination of interaural time and intensity in the lateralization of high-frequency complex signals," *J Acoust Soc Am*, vol. 87, pp. 1702-8, Apr 1990.
- [8] L. R. Bernstein and C. Trahiotis, "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," *J Acoust Soc Am*, vol. 112, pp. 1026-36, Sep 2002.
- [9] N. A. Macmillan and C. D. Creelman, *Detection Theory: A User's Guide (2nd edition)*. Mahwah, NJ: Lawrence Erlbaum Associates, 2005.
- [10] S. Furukawa, "Detection of simultaneous modulation of interaural time and level differences: Effects of modulation rate and relative phase (L)," *J Acoust Soc Am*, vol. 132, pp. 1-4, Jul 2012.
- [11] D. W. Grantham, "Spatial hearing and related phenomena," in *Hearing*, B. C. J. Moore, Ed., ed San Diego: Academic Press, 1995, pp. 297-345.
- [12] S. Furukawa, *et al.*, "How independent are the pitch and interaural-time-difference mechanisms that rely on temporal fine structure information?," in *16th International Symposium on Hearing (ISH2012)*, Cambridge, UK, 2012.
- [13] B. A. Wright and M. B. Fitzgerald, "Different patterns of human discrimination learning for two interaural cues to sound-source location," *Proc Natl Acad Sci U S A*, vol. 98, pp. 12307-12, Oct 9 2001.
- [14] G. A. Moore and B. C. Moore, "Perception of the low pitch of frequency-shifted complexes," *J Acoust Soc Am*, vol. 113, pp. 977-85, Feb 2003.