

## スペクトル変形同定の聴覚トレーニングにおける適応的フィードバックの影響\*

☆加庭輝明(筑波大学), 金成英(ヤマハ), 寺澤洋子(筑波大学), 伊藤寿浩(ヤマハ),  
池田雅弘(ヤマハ), 山田武志(筑波大学), 牧野昭二(筑波大学)

### 1 はじめに

音響に関連した業務に携わるエンジニア, 特にレコーディングエンジニアにとって聴覚能力の向上は重要である。なぜならエンジニアは最終的にリスナーが耳にする音源の品質に関する責任を負うからである。Miskiewiczは「...the recording engineer must know what technical parameters of the acoustic signal have to be modified in order to create the desired quality of sound.」と述べている[1]。日本では, このような聴覚能力を開発するため, グループレッスンと講義を主体とした「聴能形成」が様々な大学および企業で継続的に行われており[2], 2011年1月6日に行われた音響教育研究会では最新の実施例が多数紹介された[3-8]。

また, カナダのマギル大学を中心としたグループでは, コンピュータを使った個別トレーニングシステムの開発が進められている[9, 10]。このようなシステムでは, インストラクターが訓練生の成績を確認しながら適切な指導を行うためにコンピュータが使用され, エンジニアの業務経験を通じた識別能力や同定能力の向上に比べ, より早く効率的な能力向上が可能である。ただし, 適切なアドバイスや指導を提供し, 訓練生のために練習メニューを作成するインストラクターが不可欠である。

そこで, 本研究では, 学習履歴データベースを導入し, 訓練生の成績履歴に基づいて苦手な部分を算出し, 重点的に練習させる適応的な独習用の聴覚トレーニングシステムを提案する。システムが自動でデータベースを参照し, 適応的な処理を施すことでインストラクターの働きを担うことになり, 効率の良い独習が期待できる。また, このような適応的

フィードバックを与えた場合にスペクトル変形の識別能力の向上にどのような変化が見られるか, 評価実験および検討を行った。

### 2 聴覚トレーニング

聴能トレーニングの分野では上記で述べた通り, 様々な種類のトレーニングプログラムが提案されている。マギル大学の研究グループでは(1) Matching, (2) Removing, (3) Identificationといった3種類のトレーニングが当初提案された[9]。どのトレーニングにおいても訓練生は操作によるリファレンスを聴きながらトレーニングを行うことが出来る。一方, 「Absolute Identification」と呼ばれる, 訓練生がリファレンスを参照しないで変形されたスペクトルの同定を行うトレーニングも存在し[9, 10], 本研究ではこの方式を採用した。具体的には, ある中心周波数・dB値・Q値を持ったフィルタに通したサウンドとフラットな状態のサウンドとを訓練生に聴き比べてもらい, スペクトルが変化した中心周波数が何Hzであるかをユーザーが回答するシステムとなっている。

### 3 適応的フィードバック

周波数帯域の同定において, 人によって苦手とする帯域が異なる。そこで本研究では個人の学習履歴をデータベースに保存し, 周波数帯域ごとの正答率によって問題を選択するランダム関数に適量の重みを与え, 各周波数帯域の出現確率を相対的に調整するという手法を提案する。つまり, 各周波数帯域に対する重みを全体の重みの総和で割ることで, 実際の出現確率が算出される。そうすることで, 正答率の高い帯域に比べ, 正答率の低い帯域の出現率が増加し, 苦手とする帯域をより効果的かつ効率的にトレーニング出来ると考えられる。また, トレーニングを行う中で苦手

\* Investigating the influence of adaptive feedback on the ear training in identifying spectral modifications, by Teruaki KANIWA(University of Tsukuba), Sungyoung KIM(YAMAHA), Hiroko TERASAWA(Univ. of Tsukuba), Toshihiro ITO, Masahiro IKEDA(YAMAHA), Takeshi YAMADA, and Shoji MAKINO(Univ. of Tsukuba)

とする帯域の正答率が上がると、システムが出現率を自動的に低下させ、個人の能力に適應したトレーニングが可能となる。図1に本研究で提案するトレーニングシステムのブロック図を示す。

各帯域の刺激に対する相対的な出現確率を式(1)に示す。

$$W_N = 100 - \frac{100 - L}{100} \cdot R_N \quad (1)$$

ここで、 $W_N$ は帯域 N に属する刺激の相対的な出現確率を、 $L$ は最低出現確率を、 $R_N$ は正答率をそれぞれ表す。出現確率  $W_N$ は正答率  $R_N$ の増加に伴って線形に単調減少する形となっている。また、最低出現確率  $L$ を25に設定し、相対的な出現確率  $W_N$ が25~100の範囲で値をとるようにした。この相対的な出現確率を周波数ごとに MAX/MSP 上で用意されたランダム関数用の iTable object に与えることにより、重み付けされた出現確率を実現している。提案法では、苦手とする帯域をより多く練習することが出来る適應的なシステムである。(以後、Smart Random と呼ぶ。)

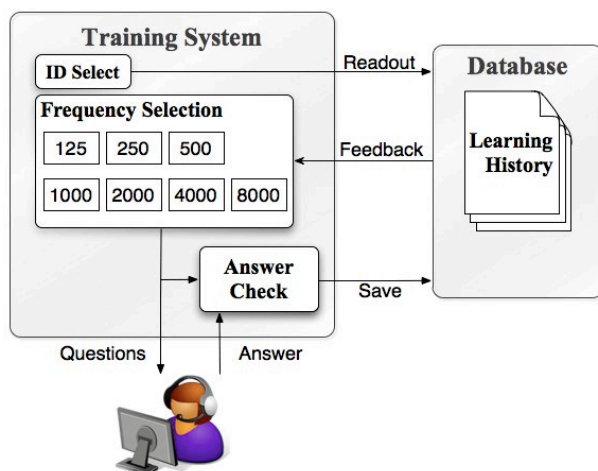


Fig. 1 システムのブロック図

#### 4 実験装置・方法

実験場所にはマルチメディアスタジオを使用した。また、実験に用いるサウンドファイルにはピンクノイズ、オーケストラ、ピアノの3つを用意した。これらのサウンドファイルはトレーニングで対象とする周波数帯域に成分を持つように作成した。その他実験に使用した機器を表1に、またトレーニングプログラムの実行画面を図2に示す。実験はスタジオにおいて週に2日30分程度の時間で個別に行い、表2に示す内容のトレーニングを1

ヶ月間行った。トレーニングで行う課題は週ごとに難易度が上がる構成となっている。

Table 1 使用機器

OS	MAC OS X ver.10.6.4
Audio Interface	MOTU Ultra Lite
ヘッドフォン	SENNHEISER HD 650
プログラム言語	Max MSP ver.5.1.5

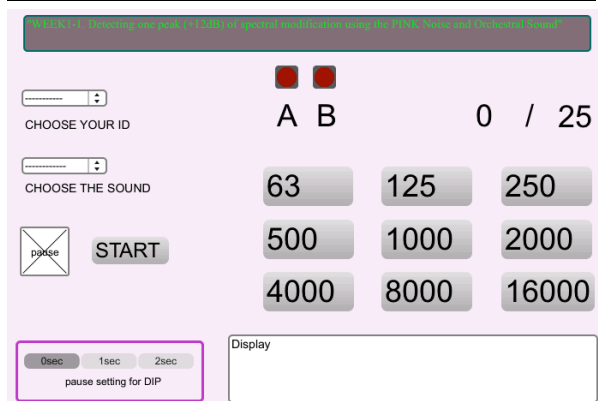


Fig. 2 トレーニング実行画面

Table 2 トレーニング内容

週	dB 値	Sound File
1	+12 dB, +6 dB	Pink Noise, Orchestra, Piano
2	+6 dB, +3dB	Pink Noise, Orchestra, Piano
3	-12 dB	Pink Noise, Orchestra
4	±12 dB, -12 dB	Pink Noise, Orchestra

評価実験は8人の被験者に対して、製作したシステムで実際にトレーニングを行った。適應的フィードバックの影響を調べるため、ランダムに問題を出题する従来システム (Normal グループ) と学習履歴データベースを用いたシステム (Smart グループ) の2グループに被験者を分けて実験を行った。グループを分ける際には25問のテスト (+12 dB, Pink Noise) を行い、正答率の平均値・標準偏差が均等になるようにグループを分けた (表3)。

Table 3 被験者のグループ化

	Normal	Smart
Mean [%]	64	63
Std-Dev.	4.89	5.19

#### 5 実験結果

評価実験の結果を以下に示す。図3, 4, 5にピンクノイズ、オーケストラ、およびピアノに対する正答率の平均値および標準偏差を週ごとにそれぞれ示す。図3, 4, 5を比較すると、Smart グループの正答率の平均値がどの

サウンドでも高くなっている。また、これらの結果をサウンドごとに平均をとり、グループ別にまとめたデータを図6に示す。

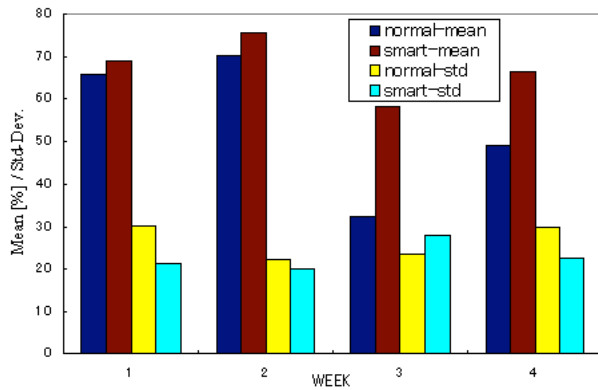


Fig.3 ピンクノイズにおける週ごとの正答率の平均値と標準偏差

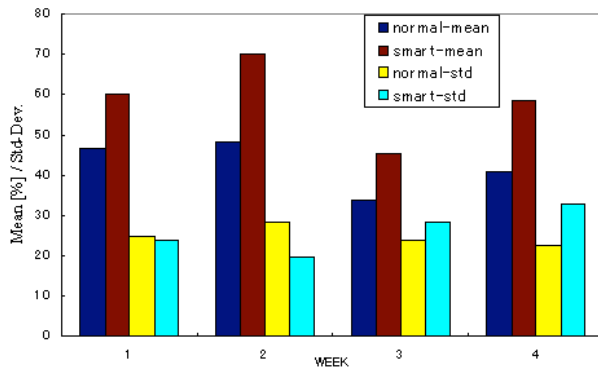


Fig.4 オーケストラにおける週ごとの正答率の平均値と標準偏差

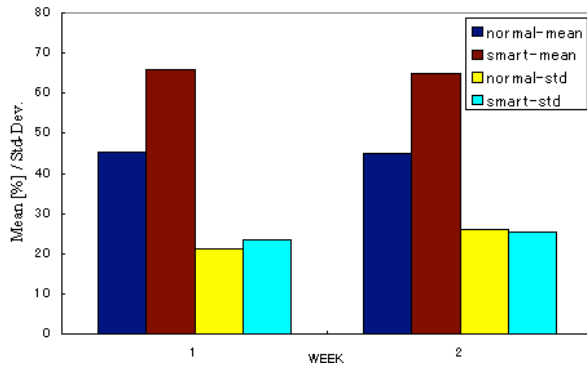


Fig.5 ピアノにおける週ごとの正答率の平均値と標準偏差

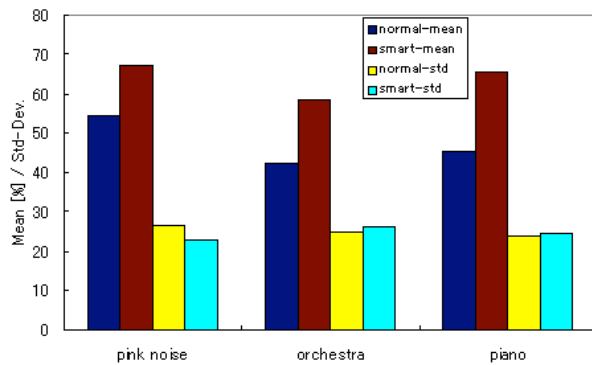


Fig.6 サウンドごとの平均値と標準偏差

NormalグループとSmartグループの結果を比較すると、図6からSmartグループの方が正答率全体の平均値が高くなっている。また、正答率のばらつきを示す標準偏差を比較すると、双方のグループに大きな違いは見受けられない。よって、グループによる個人の能力差にそれほど違いはないと言え、これらのデータは提案手法に正答率を上げる効果があることを示唆している。

今回は被験者の人数が少なく、統計的検定とはずれ値の除外は実施していないが、全体的な平均をとることに関する疑問点としては(1)得意な帯域のデータもあわせて平均を取って比較することになるので、苦手な帯域の正答率が改善していたのかどうかわかりにくい、そして(2)全体的な平均は、得意な帯域の高い正答率に影響を受けやすく、苦手な部分の低い正答率が反映されにくい、の二点が挙げられる。

そこで、苦手な帯域のみに着目し、実際の数値ではなく「正答率の変化の方向」の指標(符号の判定)を利用し、「苦手部分の改善」に関する検討も行うこととした。

具体的には、それぞれの被験者において、週ごとに正答率が低い周波数帯域を3つ抽出し、次の週にその帯域での正答率がどのように変化しているかを解析した。変化の指標は、「- (低下)」「0 (変化なし)」「+ (向上)」の3つで表すことにした。全被験者のデータをグループ別にまとめた結果を図7に示す。

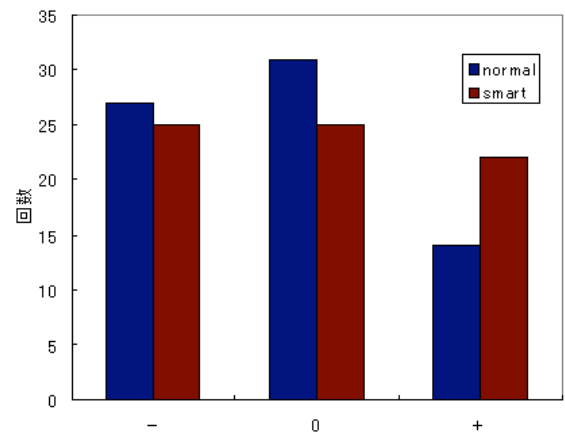


Fig.7 正答率が低い周波数帯域の、翌週における成績の変化

図7に示した変化の様子を比べると、Smartグループの方が「-」要素が少なく、また「+」

要素が多くなっている。つまり、Smart グループの方が苦手とする帯域での成績の向上が多いと言える。ここで「-」と「+」要素数の差を表すために「-」要素を -1, 「+」要素を +1 として合計をとると、その合計値は Normal グループが -13 となったのに対し、Smart グループでは -3 となり、成績が低下する割合が Smart グループの方が低くなっていると言える。

また、実際に正答率がどれ程の大きさで変化したのか、その頻度を表す度数分布を図 8, 9 に示す。図 8, 9 を比較すると、Smart グループの方が Normal グループに比べ主に「0~30%」の間で正答率の向上する頻度が多くなっている。また、「-20~-10%」も Smart グループで低くなっており、正答率が低下する頻度が少ない。これは元々成績の悪かった周波数帯域に絞って解析したものであるため、Smart グループは当初の目的である「苦手な周波数帯域の効率よいトレーニング」が行えたと言える。

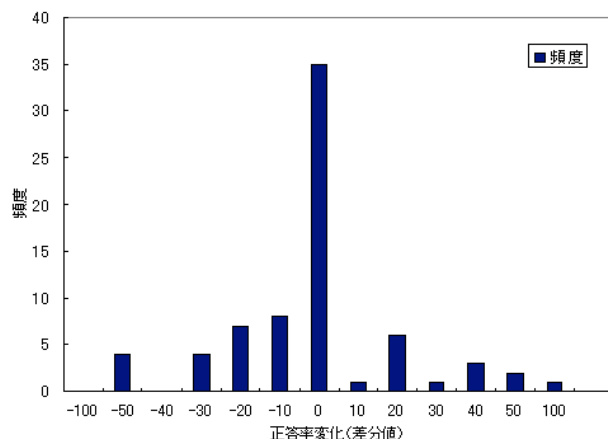


Fig.8 正答率が低い周波数帯域における正答率変化(差分値)のヒストグラム (Normal)

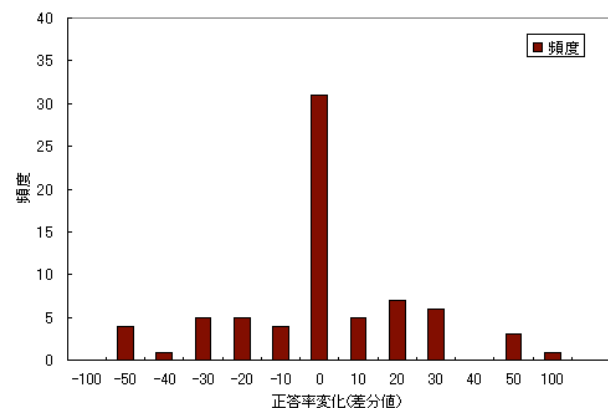


Fig.9 正答率が低い周波数帯域における正答率変化(差分値)のヒストグラム (Smart)

## 6 おわりに

本研究の提案手法である学習履歴を用いた適応的フィードバックシステムは以上の議論より従来のシステムに比べ、苦手部分の改善において有効的な結果が得られていると判断できる。今後は被験者数を増やした更なる評価実験を行い、統計的検定を行った上で、提案手法の有効性を示すことが課題となる。

## 参考文献

- [1] A.Miskiewicz, *Timble Solfege: A Course in Technical Listening for Sound Engineers*, J.Audio Eng.Soc., Vol.40, No.7/8, 621-625, 1992.
- [2] 岩宮ほか, 音の感性を育てる- 聴能形成の理論と実際, 音楽之友社 1996 年
- [3] 岩宮ほか, 聴能形成-音に関わるプロフェッショナルのための「音の感性」養成カリキュラム-, 音響教育研究会資料, Vol.6, No.1, pp.1-6, 2011 年
- [4] 丸井ほか, 東京藝術大学音楽環境創造科における聴能形成について, 音響教育研究会資料, Vol.6, No.1, pp.7-12, 2011 年
- [5] 西村明, 文科系学生に対する聴能形成の意義と成果, 音響教育研究会資料, Vol.6, No.1, pp.13-18, 2011 年
- [6] 河原一彦, 聴能形成のカリキュラムや運用について, 音響教育研究会資料, Vol.6, No.1, pp.19-22, 2011 年
- [7] 森尾ほか, 聴能形成導入の事例紹介およびトレーニング効果の持続性について, 音響教育研究会資料, Vol.6, No.1, pp.23-26, 2011 年
- [8] 伊藤ほか, ヤマハにおける聴能形成の社内展開, 音響教育研究会資料, Vol.6, No.1, pp.27-32, 2011 年
- [9] R. Quesnel et al., *Timbral Ear Trainer: Adaptive, Interactive Training of Listening Skills for Evaluation of Timbre*, Proc. Int. Conv. of AES, Preprint 4241, 1996.
- [10] J. Corey, *Audio Production and Critical Listening*, Focal Press, 2010.