

## 科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成 25 年 5 月 31 日現在

機関番号：24601  
 研究種目：若手研究(B)  
 研究期間：2011～2012  
 課題番号：23730715  
 研究課題名(和文) 時系列の聴覚表象に頑健性を与える知覚属性間の交互作用に関する心理物理的研究  
 研究課題名(英文) Psychophysical research about the interaction between perception aspects providing robustness to sequential auditory representation  
 研究代表者  
 松井 淑恵 (MATSUI TOSHIE)  
 奈良県立医科大学・耳鼻咽喉科・頭頸部外科学講座  
 研究者番号：10510034

### 研究成果の概要（和文）：

最新の聴覚モデルでは、聴覚は到来した音響信号を、蝸牛基底膜のフィルタバンクによる周波数分析の結果であるトノトピック情報と、聴神経の位相固定による周期の表現である時間間隔情報に符号化する。先行研究では、両者が統合されている可能性が示されてきた。両者が2次元空間表象として統合されると仮定し、この空間における遷移パタンの同定課題を行ったところ、仮定を支持する結果は得られなかった。ただし、聴覚初期過程の計算モデル AIM の出力では、波形の周期を表現する「安定化聴覚イメージ (Summary SAI)」にもトノトピック情報の手がかりのパターンが現れた。時間間隔情報上でトノトピック情報の手がかりと時間間隔情報の手がかりが干渉する可能性が示唆された。

### 研究成果の概要（英文）：

It has been modeled that the auditory system encodes acoustic signals into two independent informations. One is tonotopic information reflecting the frequency response characteristics of the basilar membrane, and another is periodicity information reflecting the temporal patterns of phase-locked auditory nerve firing. Based on the previous study about size perception of sound source, a hypothesis can be proposed that the tonotopic and periodicity information are combined into an internal “two-dimensional representational plane.” The current study tested this hypothesis by conducting an experiment to recognize the transition pattern of vowel-like sounds on the plane. Listeners were able to recognize the transition patterns on the axes of tonotopic and periodicity information exclusively. However, they were unable to follow the trajectories of sounds in the two-dimensional representational plane. The vowel-like sounds were simulated by a computational model of auditory system, Auditory Image Model. While the “excitation pattern” displayed transition patterns corresponding to the manipulation of resonant scaling (RS) only, “Summary SAI” displayed both of patterns of fundamental frequency (F0) and RS. It suggests the interaction between RS and F0 on the periodicity information dimension.

### 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
平成 23 年度	1,100,000	330,000	1,430,000
平成 24 年度	500,000	150,000	650,000
合計	1,600,000	480,000	2,080,000

研究分野：社会科学

科研費の分科・細目：心理学・実験心理学

キーワード：知覚、聴覚、ピッチ、サイズ、時系列パターン

### 1. 研究開始当初の背景

我々が音声を発話するとき、おなじイントネーションを意図していても、毎回同じピッチで発声できるわけではない。また同じ曲であっても常に同じピッチで歌唱が開始されるとは限らない。それでも音声のイントネーションや、音楽のメロディやハーモニーが再認されるのは、イントネーションやメロディの同一性が、ピッチの上下や音程などを含んだピッチの遷移（移り変わり）パターンであるからだと考えられる。異なるピッチで再現されたときに再認可能かどうかは、ピッチの遷移パターンの頑健性に依存する。

最新の聴覚モデルでは、我々の聴覚は、音声や音楽などの音響信号を、1) 音響信号の周期性と、2) 音響信号のスペクトル包絡の2種類に符号化しており、これらは独立しているとされている。1) は音のピッチ、2) は音色の知覚に関係する。イントネーションやメロディのピッチ系列の聴覚表象は、主に信号の周期性から構成されていると考えられる。

しかしながら、これら2種類の符号が知覚判断において影響し合うこともまた報告されている。たとえば、基本周波数の低い音のスペクトル重心が低く、基本周波数の高い音のスペクトル重心が高く設定された場合、個々の音のピッチは変わらないにもかかわらず両者の間の音程がより広がって知覚される現象が報告されている[1]。音のスペクトル重心と基本周波数が相関して変化するとき、変化のパターンに対するカテゴリー判断の正確さとスピードが向上するという報告もある [2]。

ところで、音響信号のスペクトル包絡には理論的には、音源の共鳴体（声道管、楽器のボディ）の「形状」と「サイズ」の情報が含まれている。この両者は聴覚系で自動的に分離される、とする説が近年計算モデル化され [3]、そのモデルの妥当性が実証されつつある [4]。たとえば、我々が大人と子どもでも、同じ「あ」の母音を「あ」であると聞き取ることができるのは、聴覚系による形状とサイズを分離する機能によると考えられる。

音程判断が音色に左右された実験結果[1, 2]も、このような聴覚系の機能によって、スペクトル重心が音源の共鳴体のサイズとみなされ、サイズ系列の聴覚表象とピッチ系列の聴覚表象が統合されて、「音程」の判断が行われていた可能性がある。インパルス応答から共鳴体のサイズと形状を抽出する聴覚系の機能が、対象を音声に限定しないことは実証されている[5]。

これらの先行研究から、聴覚末梢において独立に扱われる「音響波形の周期性（ピッチ）」と「音響信号のスペクトル重心（音源の共鳴体のサイズ）」が何らかの形で統合され、より頑健な時系列聴覚表象を与えるとす

る仮説を立てることができる。

## 2. 研究の目的

音のスペクトル重心と基本周波数が相関すると、音程判断の反応が速くなることを確認した研究[2]や、音声を発話しているヒトのサイズの判断に音声のピッチが影響を及ぼすことを報告した研究[4]から、ピッチとサイズが相関した音に我々の聴覚が強く適応していることがわかる。これらの先行研究は、ピッチとサイズが相関している軸に対して我々の聴覚がチューニングされている可能性を示している。我々の環境において、音源の振動数と、共鳴体のサイズが相関する場合が非常に多いために（例：純音、ピッチをもつ打楽器など）、我々の聴覚系がそのような音響信号の連続の処理に対して最適化されることは合理的である。この仮説が正しければ、ピッチとサイズが相関する場合の遷移の音系列はより表象が明瞭になるはずである。

また、ピッチと共鳴体のサイズの相関が我々の聴覚系にとって特殊な関係であることを検証するには、両者がどのような関係で統合されているかをまず確認しなければならない。統合にはさまざまなレベルでの統合が考えられるが、より初期の段階の統合として考えられるのは、両者が2次元平面上を構成する統合方法である。聴覚末梢の計算モデルである Auditory Image Model[6]でも、このような統合は間接的に示唆されている。この仮説を、2次元平面上の遷移パターンを追跡できるかどうかを確認する心理物理的実験で検証した。

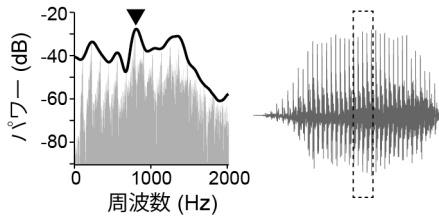
## 3. 研究の方法

実験には聴力に問題のない京都市立芸術大学の学生3名が実験に参加した。

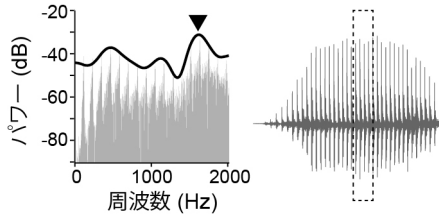
実験刺激の素材には成人男性による母音“a”の発話音声を使用した。48 kHz サンプリング、持続時間は406 ms、基本周波数は定常部で約112 Hzである。STRAIGHT[7]を用い、オリジナルの母音のスペクトル包絡と、基本周波数を変化させた母音を複数作成した。図1はSTRAIGHTでスペクトル包絡と基本周波数を変化させたときの波形と、特定の1フレームのスペクトルを示す。

実験刺激は、図2の1つの円上にある12の母音を、円周の順番に時間軸上に並べた音列である。音列の最初の母音（図3のVI, IX）と、円周の回転の方向（時計回り・反時計回り）の2項目の組み合わせが異なる音列を、円ひとつあたり4種類、計20種類用意した。個々の母音間の空白時間は100 ms、音列の持続時間は6056 msである（図3）。

(A) オリジナルの母音の波形と、STRAIGHT 分析合成の 1 フレーム分のスペクトル



(B) スペクトル包絡の伸縮率を 2 倍、波形の周期を 1 倍で変換した波形と、STRAIGHT 分析合成の 1 フレーム分のスペクトル



(C) スペクトル包絡の伸縮率を 2 倍、波形の周期を 2 倍に変換した波形と STRAIGHT 分析合成の 1 フレーム分のスペクトル

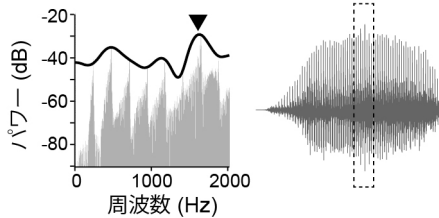


図 1 母音のスペクトル包絡と基本周波数の変換による波形 (右) とその一部のスペクトル (左)。下向きの三角形はフォルマントの位置が操作によって変化することを示す。破線はスペクトルを示した範囲。

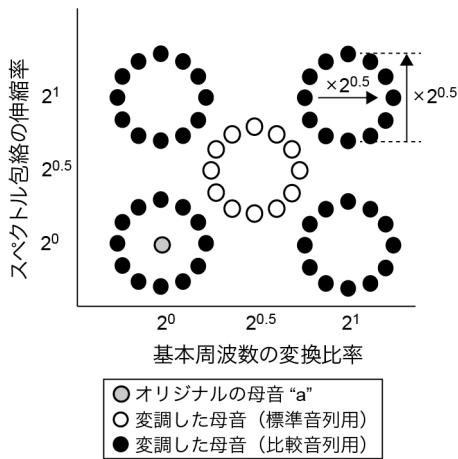


図 2 刺激音列を構成する母音の、オリジナル母音に対するスペクトル包絡の伸縮率と基本周波数の変換率。

参加者が同定する対象は、(A) スペクトル包絡の伸縮率と基本周波数の変化率による仮想的な 2 次元空間における回転方向、(B)

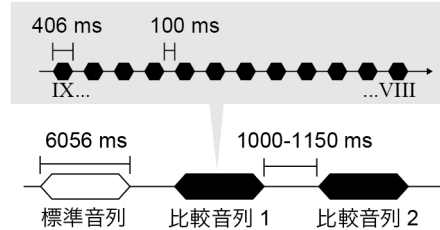
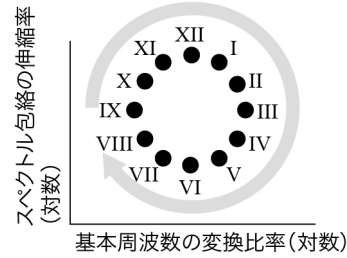


図 3 図 2 に示した個々の母音を時系列に並べ、音列を作成する方法。

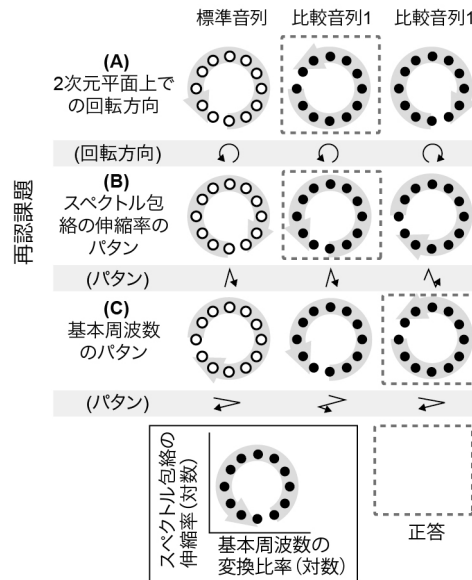


図 4 課題ごとの標準音列および比較音列の例と、各ケースの正答例。

スペクトル包絡の伸縮率の遷移パタン、(C) 基本周波数の変化率の遷移パタンの 3 つである。参加者への説明では、スペクトル包絡の伸縮率は「話者のサイズ」、基本周波数の変化率は「母音のピッチ」とした。以上の課題は、標準音列、比較刺激 1、比較刺激 2 の XAB 法で実施した。刺激間間隔は 1000 ms から 1150 ms の間でランダムに変化させた。刺激の例正答例を図 4 に示す。

課題(A)、(B)、(C) に各 48 試行、合計 144 通りを 1 セットとし、各参加者に 1 セット 2 回ずつ実施した。平均呈示音圧は約 70 dB SPL (音列ごとに  $\pm 10$  dB の範囲で変動) とした。課題は計算機 (Apple iMac) 上で動作する DSP システム Kyma (Symbolic Sound) で呈示し、同じく Kyma 上の GUI で回答させた。

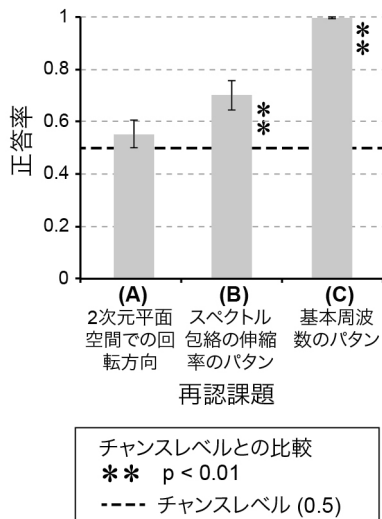


図5 同定課題の結果。3人の参加者の計6セッションをプールしている。エラーバー：標準誤差。

実験は京都市立芸大内の防音室内にて行い、刺激はヘッドフォン (Sennheiser HD600) を通して呈示した。試行ごとに正答をフィードバックした。

3つの同定課題に対する参加者ごと、セットごとの正答率の全平均を図5に示す。正答率がチャンスレベルから有意に差があった課題は(B) ( $t(5) = 3.56, p = 0.016$ ) と(C) ( $t(5) = 143.00, p < 0.0001$ ) であった。(A) の正答率にはチャンスレベルとの有意差はみられなかった ( $t(5) = 0.99, p = 0.37$ )。結果は、基本周波数の変化率の遷移パターンとスペクトル包絡の伸縮率の遷移パターンは、一方の変化に関わらず、残る一方を抽出できることを示した。ただし、基本周波数の変化率に比べて、スペクトル包絡の伸縮率の追跡が難しいことが正答率の差から窺える。スペクトル包絡の伸縮率と基本周波数の変化率による仮想的な2次元空間表象における回転方向は、同定が困難であることが示唆された。

実験結果から、トノトピック情報と時間間隔情報が2次元空間表象をとる可能性は低いと考えられる。しかしながら、時間間隔手がかりがトノトピック手がかりによる共鳴体サイズ判断に影響を及ぼすことを示した先行研究[4]から、両者が何らかの形で統合されている可能性は否定できない。そこで、聴覚の計算モデルである Auditory Image Model[6]を用い、興奮パターンと安定化聴覚イメージ (Summary SAI) における刺激の表現を観察した。図2の各円上において、同じ開始位置・同じ回転をとる音列を24 kHzにダウンサンプリングし、AIM-MAT v1.6の gm2002/dcgc/hl/sf2003/ti2003 のモジュール群によるシミュレーションを行った (図6)。

トノトピックな情報に対応する興奮パターンでは (図6、各2枚組の左パネル) スペク

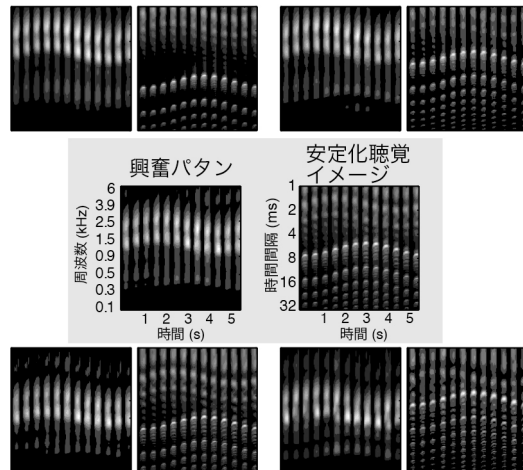


図6 標準音列および比較音列の AIM によるシミュレーション。2枚組のパネルは、図2の5つの円にそれぞれ対応する。音列の開始位置と回転方向は同一である。

トル包絡の伸縮によるフォルマントの遷移パターンのみが見られる。時間間隔ヒストグラムに対応する安定化聴覚イメージ (図6、各2枚組の右パネル) では、基本周波数の遷移パターンに加えて、短い周期の領域でスペクトル包絡伸縮率の遷移パターンと一致したパターンが見られた。

実験では、基本周波数の遷移パターンの正答率はほぼ90%を超えていた。安定化聴覚イメージからもっとも可能性の高い基本周波数の遷移パターンを抽出することは、たとえ異なる遷移パターンが異なる周期の範囲で同時に存在していたとしても、可能であると言える。一方で、スペクトル包絡伸縮率の遷移パターンは、興奮パターンにも安定化聴覚イメージにも現れているにも関わらず、同定成績は70%程度であった。話者のサイズの変化は減多にない現象であり、それを追う課題の難しさが同定成績の要因のひとつとして考えられるが、安定化聴覚イメージにおいて基本周波数の変化パターンがスペクトル包絡伸縮率パターンの判断に干渉した可能性は、今後詳細に検討する必要がある。

#### 4. 研究成果

音響信号の周期性 (時間間隔情報) と、音響信号のスペクトル包絡 (トノトピック情報) の2種類の符号化が、我々が普段聞く時系列の音の知覚にどのように関わるか、という点を明らかにするために心理物理実験を行った。この2種類の情報が何らかの形で統合されているという仮説を立て、さらに両者が2次元空間表象として統合されると仮定し、この空間における遷移パターンの同定が可能かどうかを確認することで、2次元平面上の情報として統合されているとする仮説を

検討した。実験結果からは、仮定を支持する結果は得られなかった。ただし、聴覚初期過程の計算モデル Auditory Image Model の出力では、波形の周期を表現する安定化聴覚イメージ(Summary SAI)にもトノトピック手がかりのボタンが現れた。時間間隔情報上でトノトピック手がかりと時間間隔手がかりが干渉する可能性を示す結果であるといえる。

聴覚における各符号化と、各感覚(音の高さ、音の大きさなど)との関連は、詳細が調査されてきており、ヒトの聴覚末梢がどのような働きをしているかについての情報は多くの蓄積がある。しかし、我々が普段聞く音の世界は、各感覚を別々に検討して「同じメロディだ」「嬉しそうな声だ」と判断しているわけではない。末梢で細かく符号化され、処理された音の情報が、音の知覚をどのように組み立てるか、という点に関しては、まだまだ研究の余地が残されている。

さらに、感覚・知覚という我々の内面に関する研究は、心理物理学的研究のみならず、神経科学分野にも広がりを見せている。神経科学分野における研究は、最近目覚ましい研究成果を上げているが、各種の脳機能測定方法が知覚世界まで計測するものではない以上、さまざまな工夫をして内面をはかろうとする心理物理学的研究はこの分野に対して大きく貢献できるはずである。

この研究の展望に関しては、当初の研究計画を見直し、シミュレーション結果をより活用した仮説とその検定の方向に進むほうが効率的であると考え。トノトピック情報と時間間隔情報が相関する場合と相関しない場合で、音の時系列情報の明瞭性が変化するとする仮説とその検証も、シミュレーションを利用することで、よりポイントをしぼった、シンプルな実験として計画できると考える。

#### 参考文献

- [1] Russo, F. A. & Thompson, W. F. (2005). Perception & Psychophysics, 67, 559-568.
- [2] Melara, R. D. & Marks, L. E. (1990). Perception & Psychophysics, 48, 169-178.
- [3] Irino, T. & Patterson, R. D. (2002). Speech Communication, 36, 181-203.
- [4] Smith, D. R. R. & Patterson, R. D. (2005). JASA, 118, 3177-3186.
- [5] van Dinther, R. & Patterson, R. D. (2006). JASA, 120, 2158-2176.
- [6] Patterson, R. D. et al. (1995). JASA, 98, 1879-1894.
- [7] Kawahara, H. et al. (1999). Speech Communication, 27, 187-207.

#### 5. 主な発表論文等

〔雑誌論文〕(計5件)

(1) Toshie Matsui and Minoru Tsuzaki (2013). "Interaction between resonant scaling and fundamental frequency on codings in auditory system," Proceedings of ICA2013, Montreal, Canada.

査読あり

DOI: 10.1121/1.4800162

(2) Toshie Matsui, Ryota Shimokura, Tadashi Nishimura, Hiroshi Hosoi, and Seiji Nakagawa (2013). "Speech intelligibility of hearing impaired participants in long-term training of bone-conducted ultrasonic hearing aid," Proceedings of ICA2013, Montreal, Canada.

査読あり

DOI: 10.1121/1.4799193

(3) Minoru Tsuzaki<sup>1</sup>, Chihiro Takeshima, Toshie Matsui, Toshio Irino (2013). "Perceptual outcomes by rapid alternation of the resonant scaling and its relation to the fundamental frequency," Proceedings of ICA2013, Montreal, Canada. 招待講演、査読あり

DOI: 10.1121/1.4800161

(4) Toshie Matsui, Satomi Tanaka, Koji Kazai, Minoru Tsuzaki, and Haruhiro Katayose. (2013). "Activation of the left superior temporal gyrus of musicians by music-derived sounds," NeuroReport, Vol. 24(1), 41-45. 査読あり

DOI: 10.1097/WNR.0b013e32835c1e02

(5) Minoru Tsuzaki, Chihiro Takeshima, Toshie Matsui (2012). "Pitch perception for sequences of impulse responses whose scaling alternates at every cycle," Basic Aspects of Hearing, Springer, Chap 17 (Proc of 16th International Symposium on Hearing, St. John's College, Cambridge). In press.

査読あり

〔学会発表〕(計7件)

(1) 松井淑恵、下倉良太、西村忠己、細井裕司、中川誠司。"単語理解度と語音明瞭度に対する骨導超音波補聴の効果 —最重度難聴者2名を対象としたリハビリテーション—,"日本音響学会春季研究発表会講演論文集, 658-65、東京工科大学八王子キャンパス(東京)、2013年3月14日。

(2) Minoru Tsuzaki; Chihiro Takeshima; Toshie Matsui; Toshio Irino (2013). "Matching of the Dominant Pitch of Scale Alternating Wavelet Sequences Against Complex Tones with Odd Harmonics Attenuation: Decision Statistics on

the Stabilized Auditory Images,”36th Annual MidWinter Meeting of The Association for Research in Otolaryngology, Baltimore, USA, 2013年2月16日.

(3) 松井淑恵, 津崎実. “聴覚系におけるトノトピック手がかりと時間間隔手がかりの独立性,” 日本音響学会秋季研究発表会講演論文集, 489-490. 信州大学長野キャンパス (長野)、2012年9月21日. 招待講演

(4) 津崎実, 竹島千尋, 松井淑恵, 入野俊夫. “スケール変換したインパルス応答が交替する系列に対するピッチ知覚,” 日本音響学会春季研究発表会講演論文集, 583-586. 神奈川大学横浜キャンパス, 2012年3月15日.

(5) 松井淑恵, 柳井修一, 下倉良太, 細井裕司. “反射・残響の少ない音環境における語音明瞭度評価 —残響と外部騒音による影響—,” 日本音響学会春季研究発表会講演論文集, 1223-1224. 神奈川大学横浜キャンパス, 2012年3月13日.

(6) 津崎実, 入野俊夫, 竹島千尋, 松井淑恵. “寸法知覚を中心とした聴覚情景分析 -物理世界と心理世界をつなぐ聴覚-,” 日本音響学会秋季研究発表会講演論文集, 1437-1440. 島根大学, 2011年9月20日. 招待講演.

(7) Minoru Tsuzaki, Toshie Matsui, Chihiro Takeshima and Toshio Irino. “Pitch perception for sequences of pulses alternating different resonance scales,” J. Acoust. Soc. Am., Vol. 129, No. 4, Pt. 2, 259, Seattle, Washington, USA, 2011年5月26日.

[その他]

ホームページ等

<http://www.tara.tsukuba.ac.jp/~tmatsui/>

## 6. 研究組織

### (1)研究代表者

松井 淑恵 (MATSUI TOSHIE)

研究者番号 : 10510034