

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年6月18日現在

機関番号：24301
 研究種目：基盤(B)
 研究期間：2009年度～2011年度
 課題番号：21330170
 研究課題名（和文）音声・音響信号に備わる寸法恒常性による音脈分凝と音色知覚の時間追従性
 研究課題名（英文）Temporal resolution of speech perception and auditory segregation in terms of perceptual constancy in the size provided by auditory signals
 研究代表者
 津崎 実 (TSUZAKI MINORU)
 京都市立芸術大学・音楽学部・教授
 研究者番号：60155356

研究成果の概要（和文）：音の物理的媒介物は大気圧の局所的な時間変動であるが、その発生には多くの場合何らかの音源すなわち物理的な実体が存在している。これらの物体は空間内で一定の寸法を持っており、それが安定していることが各音源に対応した音脈の形成に繋がると想定される。この研究では音源の寸法の変化に対応した共鳴スケールの時間変動が音脈分凝を中心とした音の知覚的体制化に与える影響について調べた。その結果、寸法情報の抽出がかなり高速に可能であり、その急激な時間変化によって系統的に音脈分凝ができることが分かった。

研究成果の概要（英文）：Sensation of sounds is mediated by temporal, local fluctuation of air pressure which is mostly propagated from sound sources, i.e., physical objects. The constancy of the source size is assumed to be a reliable cue for sound source segregation. The purpose of the current study was to investigate effects of the temporal change in a resonant scale corresponding the source size on auditory streaming. Results of several psychophysical experiments indicated that the resonant scale information could be extracted in a fast process, and that its rapid alternation could cause stream segregation in a systematic manner.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	8,500,000	2,550,000	11,050,000
2010年度	3,200,000	960,000	4,160,000
2011年度	2,900,000	870,000	3,770,000
総計	14,600,000	4,380,000	18,980,000

研究分野：心理学

科研費の分科・細目：実験心理学

キーワード：声道長正規化、寸法情報知覚、聴覚、聴覚による情景分析、聴覚系の時間分解能、聴覚系の時間積分、音声の個人性知覚、音脈分凝

1. 研究開始当初の背景

(1) 人間の音声には、例えば成長に伴う身体の大きさの変化や、成熟に伴う男女間の身体

構造の差が存在している。これらの差異は声道の寸法に影響を与え、同一の母音であっても母音識別の手掛かりとなる共振周波数は

異なったものとなっている。しかし人間は寸法に依存しない母音の知覚が可能である。この事実から寸法の正規化を行う聴覚的計算過程が存在しているとする計算理論が提案されていた。この理論は学習を重視する伝統的な音声正規化研究の立場と一線を画しており、この生得的な寸法正規化のための寸法知覚が実際に人間の聴覚系で実施されていることを証拠づける知覚実験結果も報告されていた。この新規性の高い領域については、ピッチやラウドネスに比べて知覚特性を調べた研究も少なく、特に時間的変化に対する追従性はほとんど調べられていない状況であった。

(2) この共振特性については特に音声と重要な知覚対象に着目すると、性差や体格などの音源の違い、すなわち個人性の違いを表す特性となりうる。音源情報は聴覚情景分析という視点からも重要な要因であるが、この聴覚情景分析の研究は1990年以降に急速に関心が高まった一方で、それは知覚心理学の研究としてはまだ歴史も浅く未解明な部分も多く存在している。さらに音声生成の観点からは、個人性は発声者の声道の寸法だけで媒介されるとは決まっておらず、他の要因についても検討することの必要性が高かった。

2. 研究の目的

本研究の目的は、聴覚的に与えられる音声の寸法情報の恒常性を利用して、人間の聴覚系がどこまでの確に音源の分離（音脈分凝）ができるかを解明することであった。ものの寸法という概念は視覚情報から主に与えられると考えがちであるが、実際には共鳴体の共鳴周波数にも寸法情報は系統的に反映されている。これは心理学的にはあまり関心が寄せられていなかった側面である。研究の目的を達するために、以下の3点に着眼した実験的研究の実施を計画した。(a)寸法変化の速さと変化幅に対する音脈分凝の成立範囲の確定。(b)寸法変化に対する聴覚的時間追従性の測定。(c)音声の個人性知覚における寸法とその他の要因との分離とその機能差の検討。

3. 研究の方法

(1) (a)について調べるために、2種類の音A, BをABA-ABA-の反復で繰り返すギャロッピング・パターンを作成して探求した。AB音の間の知覚的距離が広がるほどギャロッピング・リズムを聴くことが困難となり、AとBの独立した2つの音脈が聞こえてくるのが従来から知られている。この研究の狙いは共鳴体の寸法の違いによって音脈分凝が生じるか否かを確かめることであるが、単にある音の共鳴周波数をスケールしただけでは寸法という概念を持ち出さずに、より単純な音色の違いによって音脈分凝が生じたに過ぎないという仮説が棄却できない。この仮説を排除すべく、AB音間にはすべての場合母音範疇の違い（これは、声道形状の違いに起因する音の違いとなる）を施した。

(2) 上記のパラダイムは声道の寸法要員以外に声帯の駆動周波数に対しても導入することが可能である。この両者は物理的な空間では独立に制御可能な変数であり、聴覚初期過程においても2つの異なる表現上の違いになりうるものである。両者が知覚情報統合上でも完全に独立か否かを調べるために、両変数の変化方向の相関を要因としてキャッピング・パターンを作成し実験を行った。

(3) (b)の寸法変化に対する時間的な追従性を調べるための基礎データとして寸法弁別閾の刺激持続長への依存性を調べる実験を実施した。

(4) (3)の研究の過程で新たに問題意識に上がったのは持続長という概念そのものであった。通常、母音や楽音など明確なピッチを有する音は調波複合音と言われる持続的かつ周期的構造を持つ音である。それに対して共鳴体などにパルスを一とつ与えた（叩いた）音はインパルス応答と言われる過渡音であり、持続時間の定義になじまない。母音などは実は複数のインパルス応答が時間軸上で周期的に重畳加算された音である。よって最も急速な共鳴寸法の交替は各員パルス応答毎に生じるのである。このような刺激に対してどのような知覚反応が生じるかを調べるために、スケール交替ウェーブレット系列（Scale Alternated Wavelet Sequence; SAWS）を作成して聴取実験を実施した。

4. 研究成果

(1)については、共鳴スケーリングの違いによって基本周波数と同様に音脈分凝が生じるという証拠を得た（図1）。さらに分凝するのに必要な最小の違いは基本周波数よりも共鳴スケーリングの方が小さい傾向にあるという結果を得た。共鳴スケーリングの違いに

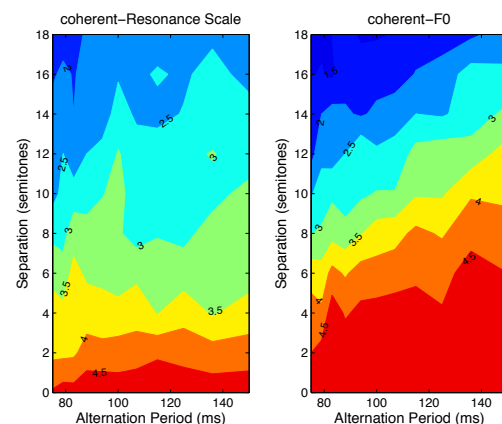


図1 ギャロッピング・リズム・パラダイムによる共鳴スケーリングと基本周波数の差による音脈分凝。左パネルが共鳴スケーリング、右パネルが基本周波数によるもの。各パネル、横軸は交替時間、縦軸は両指標間の距離であり、寒色系であるほど音脈分凝が起こっていることを示す。

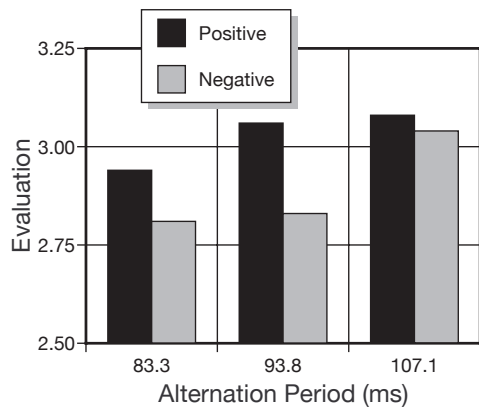


図2 ギャロッピング・リズム・パラダイムによる共鳴スケールと基本周波数の組み合わせ方向による音脈分凝。縦軸は音脈分凝の程度に対する知覚評価値であり、低いほど音脈分凝が生じやすいことを示す。横軸は交替周期。棒の色の違いが組み合わせの方向の違いを表す。

対する弁別閾は基本周波数の弁別閾よりも大きいため、この違いは単純な知覚的な違いの分かりやすさでは説明しにくい。音声生成において基本周波数（声帯の振動数）は自由度の高い制御が可能であるに対して、制御の自由度が低い声道長（実際には喉頭の位置）の変化の逸脱に対して分凝を起こしやすく知覚処理過程が適応していることを示唆する。

(2)については、声道寸法要因と基本周波数要因の間が負の相関を持って交替変化する場合の方が正の相関の場合よりも分凝が起こりやすいという結果を得た。正の相関方向への変化は自然な音声においても生じる変化であるため、(1)の結果と同様に自然界に生じやすい変化に対しては知覚的に異質感を感じにくいような知覚特性の存在を示唆するものである。（図2参照）

(3)については、約30msの持続長以下になると弁別閾が急激に上昇（つまり感度の急激な低下）が生じるのに対して、それ以上の持続長の増加の効果はほとんど無いことが分かった。これはピッチ知覚などによって先行研究において提唱されてきた聴覚情報処理の時間窓の限界と整合的な値であり、各種の聴覚情報処理と共通基盤で共鳴スケールに関する処理も担われていることを示す結果である。

(4)のSAWSを用いた実験では、交替するインパルス応答間のスケール変換の差がある限界を超えると各重畳加算周期の2倍の周期に対応したピッチが聞こえて来るという結果を得た。連続して到来しているインパルス応答の間に明瞭な差があればそれぞれの到来は異なるものとして処理され、その一方で2つの到来をつなげたものを基本周期として捉える知覚方略があることを示すものである。この場合のピッチシフトはいきなり1オクターブの

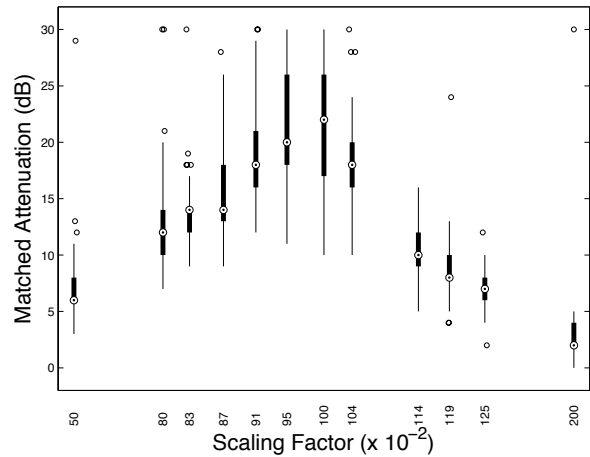


図3 SAWS刺激に対するピッチマッチングの結果。横軸はスケール変換率、縦軸はピッチマッチした奇数次高調波の減衰率。

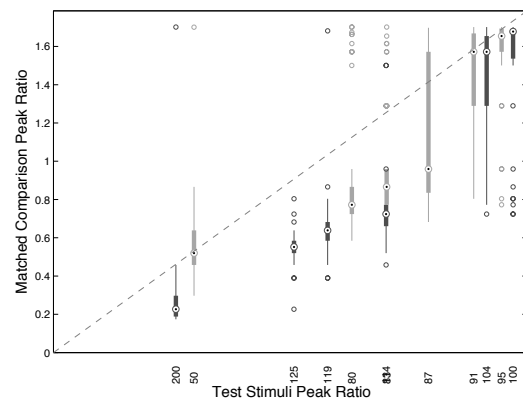


図4 図3の結果をSAWS刺激、奇数次高調波減衰刺激ともに、それぞれのスケール変換率、高調波減衰率に対するSAI上のピークの高さの比率に換算してプロットし直したものの。

飛躍があるような普段楽器音などで知覚されるピッチの変化では考えにくいものである。その一方で図3に示すように調波複合音の奇数次調波成分の減衰によって生じる変化と系統的、整合的に対応することが分かり、さらにその説明要因としては単純なスケール変換率よりも安定化聴覚像 (Stabilized Auditory Image) の上に観察される主要なピークの高さの比がよい対応を示すことが判明した（図4参照）。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計6件）

(1) Erika Okamoto, Toshio Irino, Ryuichi Nishimura and Hideki Kawahara,

- "Evaluation of voice morphing using vocal tract length normalization based on auditory filterbank" J. Signal Processing, 査読有り, 2011, 15, 283-286.
- (2) Izumida, T. & Kitamura, T., "Study of perceptual factors for speaker identification focusing on perceptual similarity of speaker characteristics" Acoustical Science and Technology, 査読有り, 2011, 32, 216-219.
- (3) Aiba, E., K. Kazai, Shimotani, T., Matsui, T., Tsuzaki, M., & Nagata, N.. "Accuracy of synchrony judgment and its relation to the auditory brainstem response: The difference between pianists and non-pianists." Journal of Advanced Computational Intelligence and Intelligent Informatics, 査読有り, 2011, 15, 962-971.
- (4) 津崎実, 徳田恵一, 河井恒, 倪晋富: "日英バイリンガル・コーパスを用いた知覚的話者空間の推定" 信学技報, 査読無し, 110. 7-12 (2011)
- (5) グリーンバーグ陽子, 加藤宏明, 津崎実, 匂坂芳典: "語彙が与える印象に基づく対話韻律合成" 日本音響学会誌, 査読有り, 67. 65-74 (2011)
- (6) Takeshima, C., Tsuzaki, M., Irino, T.: "Perception of vowel sequence with varying speaker size" Acoustical Science and Technology, 査読有り 31. 156-164 (2010)
- [学会発表] (計25件)
- (7) 竹島 千尋, 津崎 実, 入野 俊夫, "母音の持続時間が話者寸法の弁別能力に与える影響" 日本音響学会: 春季研究発表会, 2012年3月15日, 神奈川大, 神奈川
- (8) 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, 2012年3月15日, 神奈川大, 神奈川
- (9) 岡本 恵里香, 北出 晴香, 西村 竜一, 河原 英紀, 入野 俊夫, "聴覚フィルタバンクに基づく声道長推定と発話様式や身長との関係" 日本音響学会: 春季研究発表会, 2012年3月14日, 神奈川大, 神奈川
- (10) 宮森 翔子, 西村 竜一, 岡本 恵里香, 入野 俊夫, 河原 英紀, "聴覚フィルタバンクを導入した音響特徴量による若年者判別手法" 日本音響学会: 春季研究発表会, 2012年3月15日, 神奈川大, 神奈川
- (11) 宮森 翔子, 岡本 恵里香, 西村 竜一, 河原 英紀, 入野 俊夫, "若年話者判別法の音響特徴に対する聴覚フィルタバンクの導入" 情報処理学会第74回全国大会 2012, 2012年3月8日, 名古屋工大, 名古屋
- (12) Minoru Tsuzaki, Toshio Irino, Chihiro Takeshima, and Toshie Matsui, "Effects of the Correlation Between the Fundamental Frequencies and Resonance Scales as a Cue for the Auditory Stream Segregation", ARO midwinter research meeting, 2012年2月29日, San Diego, California, USA
- (13) Chihiro Takeshima, Minoru Tsuzaki, and Toshio Irino, "Discrimination of Speaker Sizes Through Speech Sounds: Dependence on Sound Duration", ARO midwinter research meeting, 2012年2月26日, San Diego, California, USA.
- (14) 岡本恵里香, 北出晴香, 西村竜一, 河原英紀, 入野俊夫 "聴覚フィルタバンクによる声道長推定と身長との相関および発話様式の影響" 日本音響学会聴覚研究会, 2012年2月4日, 那覇市IT創造館, 沖縄
- (15) 津崎実, 竹島千尋, 松井淑恵, 入野俊夫, "スケーリングした2種のインパルス応答が交替する音系列に対するピッチ知覚-調整法による心理物理実験-" 日本音響学会聴覚研究会, 2012年2月4日, 那覇市IT創造館, 沖縄
- (16) 岡本恵里香, 入野俊夫, 西村竜一, 河原英紀, "聴覚フィルタバンクに基づく声道長正規化を用いた音声モーフィングの改良" 日本音響学会関西支部, 第14回関西支部若手研究者交流研究発表会, 2011年12月18日, 産総研関西支部, 大阪
- (17) Tsuzaki, M., K. Tokuda, et al. "Estimation of Perceptual Spaces for Speaker Identities Based on the Cross-Lingual Discrimination Task" Proceedings of the 12th Annual Conference of the International Speech Communication Association, 2011年9月28日, Firenze, Italy
- (18) 岡本 恵里香, 入野 俊夫, 西村 竜一, 河原 英紀, "聴覚フィルタバンクを用いた声道長推定法の比較" 日本音響学会: 秋季研究発表会, 2011年9月21日, 島根大, 島根
- (19) 岡本恵里香, 入野俊夫, 西村竜一, 河原英紀, "聴覚フィルタバンクを用いた声道長推定" 電子情報通信学会 音声研究会, 2011年7月22日, 定山溪, 北海道
- (20) 入野 俊夫, 河原 英紀, "安定な声道長推定のための聴覚フィルタバンクとその理論" 日本音響学会: 秋季研究発表会 (招待講演), 2011年9月22日, 島根大, 島根
- (21) 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2011年秋季研究発表会（招待講演），2011年9月21日，島根大学，島根
- (22)北村達也，出水田剛志，橘亮輔，“声の類似性から個人性知覚を探る” 日本音響学会2011年秋季研究発表会（招待講演），2011年9月20日，島根大学，島根
- (23)宮森 翔子，岡本 恵里香，西村 竜一，河原 英紀，入野 俊夫 “聴覚フィルタバンクを用いた若年話者判別の検討”，日本音響学会：秋季研究発表会，2011年9月20日，島根大，島根
- (24)入野俊夫，“音声からの声道長推定における聴覚的ウェーブレット変換について”，平成23年度 数学・数理科学と諸科学・産業との連携研究ワークショップ「ウェーブレット理論と工学への応用」（主催：文部科学省・大阪教育大）（招待講演），2011年9月12日，大阪教育大，大阪
- (25)Tsuzaki, M., Matsui, T., Takechima, C. & Irino, T., “Pitch perception for sequences of pulses alternating different resonance scales.” Journal of Acoustical Society of America, 2011年5月26日, Seattle, U.S.A.
- (26)饗庭絵里子，風井浩志，下斗米貴之，田中里弥，津崎実，長国典子：“I蝸牛遅延を補正した2つのパルスの知覚的な分離と聴性脳幹反応との関係” 日本音響学会平成23年春季研究発表会，2011年3月11日，早稲田大学，東京
- (27)竹島千尋，津崎実，入野俊夫，“音話者寸法の弁別における母音の持続時間の効果：雑音駆動母音を用いた検討” 日本音響学会平成23年春季研究発表会，2011年3月11日，早稲田大学，東京
- (28)田中里弥，饗庭絵里子，津崎実，加藤宏明：“連続音響信号中の音事象検出に関わる純音の周波数遷移要因” 日本音響学会平成23年春季研究発表会．2011年3月9日，早稲田大学，東京
- (29)竹島千尋，津崎実，入野俊夫，“音の持続時間が音源の大きさ知覚に及ぼす影響：母音刺激を用いた検討” 日本基礎心理学会第28回大会，2009年12月6日，日本女子大学目白キャンパス
- (30)Takechima, C., Tsuzaki, M., Irino, T., “Influences of vowel duration on speaker-size estimation and discrimination” Interspeech 2009年9月8日, Brighton, U.K.
- (31)中川祐貴，津崎実，“音脈分凝に対するピッチと寸法の交互作用の有無について” 日本音響学会2009年秋期研究発表会，2009年9月16日，日本大学郡山キャンパス

史，鈴木陽一，牧勝弘，津崎実，(共著)「聴覚モデル」，(2011) コロナ社，195-229.

6. 研究組織

(1) 研究代表者

津崎 実 (TSUZAKI MINORU)

研究者番号：60155356

京都市立芸術大学・音楽学部・教授

(2) 研究分担者

入野 俊夫 (IRINO TOSHIO)

研究者番号：20346331

和歌山大学・システム工学部・教授

北村 達也 (KITAMURA TATSUYA)

研究者番号：60293594

甲南大学・知能情報学部・准教授

〔図書〕 (計1件)

- (1)森周司，香田徹 (編)，香田徹，日比野浩，任書晃，倉智嘉久，入野俊夫，鶴木祐

