

# 平成21年度 (2009年度) 特定領域研究 (継続領域) 研究計画調書

平成 20 年 11 月 3 日  
1 版

計画研究 公募研究 終了研究領域区分	公募研究	機関・領域・研究 項目・整理番号	12601-456-A03-0002				
新規 継続区分	新規						
研究領域	領域番号	領域略称名					
	456	情報爆発 I T 基盤					
研究項目番号	A03						
計画研究のうち 調整班	-						
研究課題番号	-						
研究代表者 氏名	(フリガナ)	ミネマツ ノブアキ					
	(漢字等)	峯松 信明					
所属研究機関	東京大学						
部 局	工学(系)研究科(研究院)						
職	准教授						
研究課題名	構造不変の定理に基づく音声アフォーダンスの提案とそれに基づく音声認識系の構築						
研究経費 (千円未満の 端数は切り 捨てる)	年度	研究経費 (千円)	使用内訳(千円)				
			設備備品費	消耗品費	旅費	謝金等	その他
	平成21年度	4,700	2,200	700	1,200	300	300
	平成22年度	4,600	1,500	700	1,200	300	900
	平成23年度	0	0	0	0	0	0
	平成24年度	0	0	0	0	0	0
	平成25年度	0	0	0	0	0	0
総計	9,300	3,700	1,400	2,400	600	1,200	
(1) 関連研究分野 (細目)	細目番号	1006		(2) 関連研究分野 (細目)	細目番号	1009	
	分野	総合領域			分野	総合領域	
	分科	情報学			分科	情報学	
	細目	知覚情報処理・知能ロボティクス			細目	認知科学	
開示希望の有無	審査結果の開示を希望する						
研究代表者 連絡先	〒 113-8656 (住所) 東京都文京区本郷7-3-1 電話番号 : 03-5841-6662 Fax番号 : 03-5841-6662 Email : mine@gavo.t.u-tokyo.ac.jp						

研究組織（研究代表者、研究分担者及び連携研究者）

氏名（年齢）	所属研究機関 部局 職	現在の専門 学位 役割分担	平成21年度 研究経費 （千円）	エフオ ート （％）
研究代表者 90273333 (42) ミネマツ ノブアキ 峯松 信明	(12601) 東京大学  (885) 工学（系）研究科（研究院）  (27) 准教授	音声言語情報処理 博士（工学） 音声アフォーダンスに基づく音声認識 系の構築	4,700	20
合計 1 名		研究経費合計	4,700	

<b>研究概要</b>				
<p>(1) 研究目的等</p> <div style="border: 1px solid black; padding: 5px; margin: 5px 0;">                     特定 (公募) - 2、3 (研究目的)、6 (今回の研究計画を実施するに当たっての準備状況等)、7 (これまでに受けた研究費とその成果等) の内容を簡潔にまとめて記述してください。                 </div> <p>申請者は、話者／マイクなどの要因に不変な音声の構造的物理表象「音声アフォーダンス」を提案している。観測される事象が全て分布として記述される場合、非線形変換を含む如何なる空間写像を施しても、f-divergence は不変である。<math>N</math> 事象 (分布) 群として観測される観測量に対して、<math>N C_2</math> 個だけある f-div. を全て求めれば、それは幾何学構造を規定する。距離行列の形で表象されるこの構造表象は変換不変であり、音声メディアにおける構造表象が「音声アフォーダンス」である。</p> <p>この構造的表象は、特定領域研究「情報学」の研究の中で得られた着想であり、その後「情報爆発」でも継続して3年間に渡り、理論的先鋭／工学的検証を重ねてきた。母音系列を連続的に発声して得られる単語群や、音素バランス単語群をタスクとした孤立単語音声認識実験において、従来の音声表象では認識率が1.4%となるような状況下でも、認識率90%以上を示すことが出来、超頑健な孤立単語音声認識を実装することができた。また、本研究を通して、自閉症や失読症者の音声情報処理に対する情報論的モデルを提供することも可能となった。</p> <p>本申請は、3年間の研究成果を更に発展させることを目的として申請している。1) 多様な話者性に対する精緻なモデル化とそれに基づく頑健性の更なる向上、2) 共鳴音、非共鳴音を分割して得られる構造表象の検討、3) 従来の音声認識技術との融合と、連続単語音声認識系の構築、4) 非音声メディアにおける構造表象と、メディア普遍の運動表象の実現、などを計画している。特に4) は脳の連合野における情報処理の計算機上の実装と解釈できる方法論を提供することを目的としている。</p>				
<p>(2) 研究計画・方法</p> <div style="border: 1px solid black; padding: 5px; margin: 5px 0;">                     特定 (公募) - 4、5 (研究計画・方法) の内容を簡潔にまとめて記述してください。                 </div> <p>平成21年度の研究計画・方法について示す。</p> <p><b>「与えられた刺激群に内在する“多様性を表現する写像関数群”の設計</b> 完全なる写像不変性が引き起こす「強すぎる不変性」問題は、対象とする多様性を数学的にモデル化し、その写像群のみに不変性が成立する枠組みを求めることで解決できる。与えられた刺激群をボトムアップに観測し、そこに内在する「多様性を表現する写像関数群」を推定・設計する一般的な枠組みを構築する。</p> <p><b>「共鳴音と非共鳴音を分割した構造的音声認識の検討</b> 既に音素バランス単語音声認識においても極めて高い頑健性を実現しているが、モデル学習時と評価時に何らミスマッチが無い場合、認識率の低下が若干観測されている。共鳴音と非共鳴音 (生成のメカニズムが大きく異なる) とでは話者性による音の変形の様子が異なるため、これらを一つの枠組みで扱うよりも、異なる2ストリームとして扱い、各々において構造構築する方法が考えられる。</p> <p><b>「韻律的特徴に関する構造表象と構造的情報処理の実験的検討</b></p> <p>韻律的特徴、特に、基本周波数パターンに関して構造的な取り扱いを検討する。ピッチコントラストのみに着目すれば、それは相対音感者がメロディーを階名で書き取る作業と等価となる。このように構造表象は、基本周波数パターンの表象として応用可能であり、これを実験的に検討する。</p>				
領域略称名	情報爆発 IT 基盤	研究機関名	東京大学	研究代表者名 氏名 峯松 信明

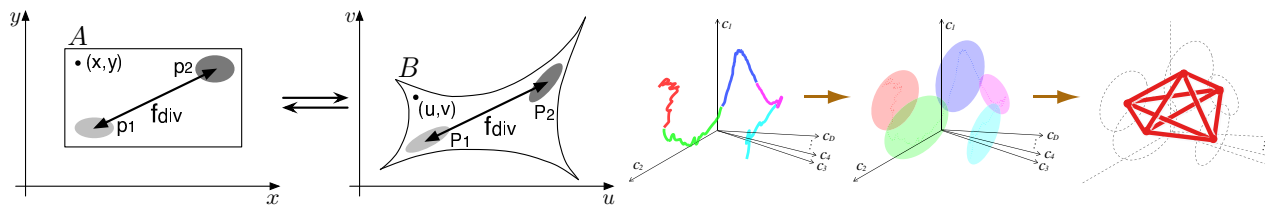
## 研究目的

本欄には、研究の全体構想及びその中で本研究の具体的な目的について、適宜文献を引用しつつ記述し、特に次の点については、焦点を絞り、具体的かつ明確に記述してください。（記述に当たっては「科学研究費補助金における評価に関する規程」（公募要領9頁を参照）を参考にしてください。）

- ① 研究の学術的背景（本研究に関連する国内・国外の研究動向及び位置づけ、応募者のこれまでの研究成果を踏まえ着想に至った経緯、これまでの研究成果を発展させる場合にはその内容等）
- ② 研究期間内に、何をどこまで明らかにしようとするのか
- ③ 当該領域の推進に貢献できる点
- ④ 応募者の専門としている研究分野と当該領域の研究が有機的に結びつくことにより新たな研究の創造が期待できる点
- ⑤ 当該分野におけるこの研究(計画)の学術的な特色・独創的な点及び予想される結果と意義
- ⑥ 平成21年度において継続して科学研究費補助金又は科学研究費補助金以外の研究費(府省・地方公共団体・研究助成法人・民間企業等からの研究費)の助成を受ける予定がある場合は、当該継続研究課題と本研究課題との相違点

①&⑤ 幼児は「大部分が母親と父親の音声、及び自らの音声」という**非常に偏った話者性の音声**の聴取を通して、誰の声（多様な声）でも対処可能になる。音声工学では従来、この多様性問題を解くために、**数万人の話者バランスのとれた音声資料**を集めてモデル化してきた。申請者は、このような「人間による音声処理」と「機械による音声処理」間の本質的差異の解消を目的として、話者／マイクなどの要因に不変な音声の構造的物理表象「音声アフォーダンス」を提案している。提案する物理表象は「線形・非線形を問わず、可逆かつ微分可能な如何なる変換に対しても一切不変である」ことを数学的に証明している（構造不変の定理）。可逆かつ微分可能な写像で対応付けられた空間 A と空間 B を考える。空間 A における種々の事象は確率密度分布として表現され、対応する空間 B の事象も確率密度分布となる。この場合、空間 A における任意の 2 分布間の f-divergence（バタチャリヤ距離、KL-div. は f-div. の一例である）は、空間 B における対応する f-div. と必ず等しい（十分性）。 $N$  事象群に対して、全ての  $({}_N C_2)$  個) 事象間距離を求めることは、その  $N$  事象群で張られる幾何学構造を規定するが、この構造が不変となる。また、「如何なる空間写像に対しても不変な量を空間中の二事象に対して求めると、それは f-div. にならざるを得ない」ことも数学的に証明している（必要性）。この不変なる構造的情報表象を音声に応用したのが「音声アフォーダンス」である。なお、事象の絶対的特性は捨て去り、事象間の関係や差異（コントラスト）に着眼する特徴抽出は、「コントラスト好き」と言われる脳の情報処理と、非常に整合性が高い。近年、人工知能・脳科学の分野で Dr.Hawkins によって提唱されている Memory Prediction Theory が注目されているが、ここでも、刺激の不変表象の重要性が説かれている（研究業績 1,2,4,5,8,9,10,11,12,15,23,26,27,29,30,31,35,36,37,38,39,41,42,43,44）。

● J. Hawkins and S. Blakeslee, *On intelligence*, Henry Holt, 2004



図：「完全変換不変量としての f-divergence」と、「f-div. を用いた運動の不変表象」

音声の不変項を導出する先行研究は、その全てが線形変換を仮定していること、「音と音の差異」の不変性ではなく「音」の不変性を議論するものである。本提案は変換の種類を問題にしないこと、また、何を不変項として考えるのか、という点において、国内外を含め、従来の研究とは全く異なる視点を導入している。音的差異への着眼は、音声を音楽として捉える方法論となるが、進化人類学では、音楽と言語の同値性が広く議論されており、本研究は、人類学研究に対して数学的枠組みを提供している。また、音声ストリームを音の列ではなく、音的差異群が成す幾何学構造として捉える枠組みは、音声を音韻列として捉える従来の言語感を否定する。しかし逆に、音声を音韻列として捉えることが困難な失読症者の情報処理モデルを与える。英語圏の 10% は失読症者であり、彼らの挙動は説明困難なパラドックスと言われていたが、「構造不変の定理」はそのパラドックスを説明できる。なお、幼児も同様、音声を音韻列として捉えることが困難であり、発達心理学研究者と共に、言語獲得の説明モデルとしての妥当性を検討している。アフォーダンス研究の第一人者である佐々木正人教授からは「音声学の地動説」と称されている（研究業績 1,2,4,5,7,9,10,16,19,26,29,30,35,36,39,40,42,43,44）

研究目的（つづき）

② 「話者、年齢、性別などの情報を音声の音響的特徴からそぎ落とすことができる」という事実は、**少数話者の音声のみを用いて不特定話者の音声認識が可能であることを意味する**。既に、音の実体の統計モデルを構築する従来法では**四千人以上の話者を用いたとしても**、例えば小人のような音声では認識率が1.4%まで下落するが、**数人の話者音声より構築した構造音声モデルでは90%以上の精度を出すことを実験的に確認している**。本研究の目的は、「音声アフォーダンス」の理論的先鋭及び工学的応用の追求である。音声の新しい「捉え方」の提案であるため、本研究の成果は、音声情報処理の全分野に対して波及する可能性が極めて高い。過去2年間の研究で、話者性に超頑健な**孤立**単語認識アルゴリズムの構築に成功している。今後2年間の工学研究としての目的は、1) 頑健性の精緻化及び、2) 従来の音声認識技術との融合、更には、3) より一般的なタスクである**連続**単語音声認識への拡張、である。また、理論研究としては、自閉症や失読症者の音声情報処理の説明モデルとしての可能性を、今まで同様、今後も追求する予定である（文献1,3,13,17,18,19,21,25,26,32,40）。

③&④ 「情報爆発」におけるヒューマンインターフェイス研究（A03班）全体に共通する問題意識は、「サイバー世界が情報爆発する以前に、物理世界は情報爆発を起こしている」事実にある。各種センサーを通して得られる多種多様な物理刺激から、意味のある情報（構造）をどのように抽出するのか。これは人工知能の分野で古くから議論されている「フレーム問題」と直結する。健常者は、**不必要な情報を「それと意識的に判断することなく」捨て去り**、効率的な情報処理を実現している。一方、「**スルーできない脳**」を持つと言われる自閉症者はこれが難しい。全てを受け取り、個別的・具体的な特徴に注意を向けてしまう。その結果、多くの重度自閉症者にとって、音声言語の運用は極端に難しくなる。母親の声しか理解できない自閉症者も存在する。このように、人間の情報処理を広く観察すると、「従来構築された音声認識アルゴリズム」を「自閉症者による音声情報処理のシミュレーション技術」として再評価することすら可能になる。申請者はこの事実を国内外の関連学会で公表、発表してきた。ここ数年、我々の主張を真摯に受け止め、理解する研究者が増えて来た。Hawkinsの主張する「不変表象」も、自閉症者の（極めて個別的・具体的な）情報処理を鑑みての主張である。

- T. Grandin 「動物感覚 ～自閉症者が読み解くアニマル・マインド～」日本放送出版協会、2006
- ニキリンコ 「スルーできない脳 ～自閉は情報の便秘～」生活書院、2008（著者は自閉症者である）

さて、音声アフォーダンスは、音声を「音+運動」として考えた時に、音声の「音」の側面を捨て去り、「運動」の側面だけを捉えた表象である。こうすることで男女を超えた音声表象が得られる。視覚野、聴覚野、体性感覚野の情報は、最終的には連合野で統合・融合される。即ち、視・聴・体という**メディアの違いを超えた、普遍の情報表象**がそこにある。「音声から音を捨てた時に、情報の普遍化が可能になった」と考えるならば、ビジョン研究から光を捨て、ロボティクス研究から身体を捨て、**全てのメディア情報を「動き」という普遍的（共通的）観点から捉え直すことで**、連合野を志向した情報処理パラダイムについて議論可能となる。Hawkinsも情報の不変表象と同時に、メディアの違いを超えた、普遍的な情報処理アルゴリズムの存在を主張し、「各領野は凡そ等価なアルゴリズムで実装されている」と述べている。既に、情報爆発関係者とこのようなブレインストーミングを開始している。また、申請者も「手の動き空間」と「音の動き空間」とを写像させることで、手を動かすことで音声生成するシステムを（構音障害者支援という目的の下）検討している。このシステムでは、**身体運動と音運動とが空間写像の関係にあるため、両者の動きは数学的・構造的には全く同一現象として観測される**。このような「もの作り」を通して、連合野の普遍的情報処理論を検討することが可能となっており、多様なセンサー入力からの「構造創発」を目的とする当該領域への貢献は甚だ大きい。

⑥ 来年度、基盤研究「音声の分節的・韻律的特徴を包含する発音の構造的表象に基づく外国語教育・学習支援」、特定領域研究「多様な目的に適した形態素解析システム用電子化辞書の開発」の支援を、継続して得る予定である。前者は発音教育支援を目的としたもので、本研究とはその性格が異なる。後者は電子化辞書整備のためのアクセント情報の整備が目的であり、本研究とは関係が無い。

領域略称名	情報爆発 IT 基盤	研究機関名	東京大学	研究代表者氏名	峯松 信明
-------	------------	-------	------	---------	-------

## 研究計画・方法

＜平成21年度の計画と22年度の計画に分けて記述してください。＞

本欄には、研究目的を達成するための具体的な研究計画・方法について、平成21年度の計画と平成22年度の計画に分けて、適宜文献を引用しつつ焦点を絞り、具体的かつ明確に記述してください。ここでは、研究が当初計画どおりに進まない時の対応など多方面からの検討状況について述べるとともに、研究計画を遂行するための研究体制について、研究代表者、研究分担者の具体的な役割（図表を用いる等）及び研究分担者とともに行う必要がある場合には、学術的観点から研究組織の必要性・妥当性及び研究目的との関連性についても述べてください。

なお、研究体制の全体像を明らかにするため、連携研究者及び研究協力者（海外共同研究者、科学研究費への応募資格を有しない企業の研究者、大学院生等（氏名、員数を記入することも可））の役割についても必要に応じて記述してください。

研究協力者として、喬宇氏（学振のポスドク）、斎藤大輔氏（東大・工学系・博士課程）、鈴木雅之氏（東大・工学系・修士課程）、桜庭京子氏（清水市福祉センター）を予定している。

過去2年間の研究において、我々は構造表象の工学応用について深い理解を得ることができた。まず、f-div. による写像不変の距離尺度であるが、これは位相幾何学において新たな位相不変量を導入し、それに基づいた工学応用を検討していることと等しい。完全なる写像不変量であり「不変性が強すぎる」問題が生じる。即ち、話者の違いを消すあまりに、単語の違いまで消えてしまう（異なる単語が同一と評価される）。この場合必要なのは、話者の違いがどのような変換関数群として実装されるのかの検討、及び、対象とする変換関数群のみに不変性が成立する音響照合方式の構築であるが、これをマルチストリーム構造化という技術を導入することで解決した。またこれとは別に、パラメータ次元数が極端に高くなる問題もあった。これについては、LDAやPCAを適切に導入することで、識別力の高い低次元空間を求めることに成功した。以上の経緯を踏まえて、2年間の研究計画を示す。

### ■平成21年度研究計画■

#### 与えられた刺激群に内在する「多様性を表現する写像関数群」の設計（担当：峯松，喬）

完全なる写像不変性のために生じる「強すぎる不変性」問題は、対象となる多様性を数学的にモデル化し、その写像群のみに不変性が成立する枠組み（照合方式）を求めることで解決できる。音声の場合、ケプストラムベクトルで考えれば、マイクの違いや、年齢・性別の違いは全て、ケプストラムベクトルに対する加算、あるいは、帯行列の掛算として近似できる。これらの変換群に対してのみ不変性を成立させる技術としてマルチストリーム構造化を提案した。しかし、音声以外のメディアへの構造表象応用を考えた場合、与えられた刺激群をボトムアップに観測し、そこに内在する「多様性を表現する写像関数群」を推定・設計する一般的な枠組みを構築する必要がある。ここでは、ガウス混合分布を用いた写像推定を拡張し、確率的な回帰写像の混合分布を考え、関数群設計を検討する。  
関連文献＝1,3,9,13,15,17,18,21,25,32,39,47（研究業績欄参照）

#### 共鳴音と非共鳴音とを分割した構造的音声認識の検討（担当：峯松，鈴木）

構造的音声認識はまず、母音連結単語（「あうえおい」など）といった人工的な語彙セットを使って検討され、本年度になって通常の音素バランス単語を用いた孤立単語音声認識が検討されるようになった。既に音素バランス単語においても、従来の技術ではおよそ実現不可能な頑健性を実現しているが、モデル学習時と評価時に何らミスマッチが無い場合、認識率の低下が若干観測されている（母音単語の場合は、このようなことは無い）。共鳴音と非共鳴音とでは話者性による音の変形の様子が異なるため、これらを一つの枠組みで扱うよりも、異なる2ストリームとして扱い、各々において構造構築する方法が考えられる。この方法以外にも、共鳴音と非共鳴音（そもそも両者は、音の生成過程が全く異なる）に対する適切な構造的処理法について実験的に検討する。  
関連文献＝1,3,9,13,15,17,18,21,25,32,39,47（研究業績欄参照）

#### 韻律的特徴に関する構造表象と構造的情報処理の実験的検討（担当：峯松，斎藤）

韻律的特徴、特に、基本周波数パターンに関して、構造的な取り扱いを検討する。母音などにおける音色の話者性は、声道長の違いによって齎されるが、基本周波数における音高の話者性は、声帯の長さ／重さの違いによって齎される。いずれも静的バイアスでしかなく、ピッチコントラストのみに

研究計画・方法（つづき）

着眼すれば、それは相対音感者がメロディーを階名で書き取る作業と等価となる。このように、構造表象は基本周波数パターンの表象として応用可能であり、この点について実験的に検討する。この時、基本周波数値を求めた後にコントラストを計算するのではなく、コントラストを直接得るための方法論を検討する。何故なら、一旦周波数値を求めようとすると、必ず誤差が混入するためである。エラーフリーの韻律的な構造抽出を検討する。

関連文献＝ 1,2,7,8,10,11,12,16,23,30,31,33,36,38,41,42,47（研究業績欄参照）

**メディア普遍かつ静的バイアス不変の情報表象を通して考える自閉症と失読症（担当：峯松，桜庭）**

提案している音声表象は、純粋に工学目的で提案された方法論であるが、自閉症や失読症者の挙動を説明できる情報処理モデルとなっている。また、健常者の場合でも、幼児の言語獲得の様子を説明できるモデルとなっていることが、これまでの研究で示されている。ここでは、このモデルの妥当性を引き続き検討し、提案する情報処理の枠組みの工学的応用のみならず、様々な分野における応用可能性について検討する。ここでは、臨床医や言語聴覚士との議論を主な活動とし、提案手法に基づく具体的な臨床手法の考案など（例えばスクリーニングテストの開発）は、直接の目的とはしない。

関連文献＝ 1,2,5,6,7,10,11,19,23,26,28,30,36,36,38,41,42,47（研究業績欄参照）

■平成 22 年度研究計画■

**従来の技術体系との融合、及び、連続単語音声認識の実装（担当：峯松，喬，齋藤，鈴木）**

音声の構造表象は、一個の音を表現することは出来ず、幾つかの音イベントの系列を、各イベント間のコントラストのみで表象する方法である。そのため、語や句を単位とした音響モデルのみ構築可能である。現在の大語彙連続音声認識技術との融合を図る場合、各仮説において単語境界で付与される言語スコアの一つとして組み込むことが考えられる（単語単位の構造スコア）。この導入によって、話者性が極めて大きく異なる入力話者に対する頑健性の飛躍的な向上が見込まれる。

関連文献＝ 1,3,9,13,15,17,18,21,25,32,39,47（研究業績欄参照）

**身体運動と音運動に対する統一的な運動表象とそれに基づく音声生成系（担当：峯松，喬）**

音声認識とは若干異なる応用であるが、「研究目的」欄に示したように、手の（身体）運動を直接、声の（音声）運動に写像するシステムを検討している。これは純粋な空間写像として実装しているので、両者の動きの構造表象を考えると、「具体的なメディアが何であるのか」という情報は完全に消え去り、手も声も「動きとしては全く同一」と見なされることになる。このようなメディア非依存・普遍的な情報処理システムを構築しつつ、ミラーニューロンとの関連性、模倣行為の創発について、検討を行なう。メディア普遍の情報処理によって初めて可能となるメディア処理を実験的に検討する。

関連文献＝ 1,9,14,15,44,47（研究業績欄参照）

**言語障害／認知障害の情報処理モデルに対する実験的検討（担当：峯松，桜庭）**

前年度の議論を引き続き継続していく。自閉症者の情報処理は、動物の情報処理と「刺激群全体に渡って内在しているパターン抽出能力が低い」という点に関して類似している。既に霊長類研究所（京都大学）の研究者との議論も開始しているが、ここでは言語進化という観点からも検討を行なう。

関連文献＝ 1,2,4,5,7,10,11,19,23,26,28,30,35,36,38,41,42,47（研究業績欄参照）

本申請課題は、情報爆発科研には初年度から参加しており、数々の成果を上げて来たことと自負している。提案している音声の不変構造に基づく音声処理系は、人間科学が見いだして来た数々の知見を通して「人間が行なう音声言語情報処理」と「機械に実装してきた音声言語情報処理」の差異を徹底的に思考して提案した方法論である。従来の音声認識処理は、様々な条件における音声を音響的に詳細にモデル化する議論を重ねて来た。申請者は、そのような情報処理論は、「刺激の具体的・詳細な側面をそのまま記憶する自閉症者（そのため、柔軟な情報処理能力が欠落する）と同じである」との主張を行なって来た。このような主張は、当初音声科学・工学者から非難されることもあったが、近年、申請者の主張を受け入れる研究者が増え、国内最大の音声工学の研究発表会である「音声言語シンポジウム」にて、10周年の特別招待講演として発表するに至っている（研究業績番号1）。

領域略称名	情報爆発 IT 基盤	研究機関名	東京大学	研究代表者氏名	峯松 信明
-------	------------	-------	------	---------	-------

## 今回の研究計画を実施するに当たっての準備状況等

本欄には、次の点について、焦点を絞り、具体的かつ明確に記述してください。

- ① 本研究を実施するために使用する研究施設・設備・研究資料等、現在の研究環境の状況
- ② 研究分担者がいる場合には、その者との連絡調整の状況など、研究着手に向けての状況（連携研究者及び研究協力者がいる場合についても必要に応じて記述してください。）

① 本研究を遂行するための研究環境の準備について、「物」的環境と「知」的環境について記す。物的環境とは音声情報処理を実行するための計算機環境、及び、研究対象となる音声を収録するスタジオや音響機器などの機材環境を指す。「知」的環境とは、情報提供者としての研究協力者や、種々の知識獲得のための図書などの情報源確保の環境を指す。

過去2年間の研究遂行によって、音声分析及び音声認識装置としての計算機環境を、一通り整備することが出来た。しかし現時点で所有している計算機群では、「刺激の多様性を表現する写像関数群の設計」の遂行に多大な時間を必要とすることが分かっており、計算機環境のより一層の充実を図る。なお、計算機の購入に際して、サーバー室などの整備が必要になる可能性が高い。こちらの方も、計算機環境向上のための経費として割当て、随時サーバー室の拡張を行なう予定である。更に、身体運動を計測するセンサーなども必要に応じて購入予定である（これらは、昨年度までの「情報爆発」科研申請に関しては、全く予算計上してこなかった物品である）。なお、必要な音響機器類や、データ収録装置の整備についても随時行なう予定である。

「知」的環境であるが、本研究の一つの特色は、工学という狭い視野に捕われず、脳科学、言語学、認知科学、言語障害学、心理学など様々な研究者との情報交換を通して、提案手法を「科学的に妥当な」方法論へと纏め上げるプロセスにあると考える。その意味において、異分野の研究者との交流は非常に大切な側面である。情報爆発科研の前身である「情報学科研」の頃より、国内外を問わず、人脈の確保については積極的に行ない、幅広い分野の研究者との「議論の場」を確立してきた。「研究目的」や「研究業績」の欄を見て頂くと分かるように、彼らとの議論が、赤ちゃん研究、言語障害研究など、従来申請者があまり参加してこなかった研究活動の場（各分野の国際会議など）において研究発表するなど、十分に結実されている。このような活動は今後も積極的に続ける予定である。なお、対話による情報収集のみならず、研究資料や図書類などを通じた情報収集も行なっており、非常に広い観点から多角的にヒト、コトバを観察することが可能になったと自負している。

② 本研究は、基本的に峯松（研究室）の単独研究であり、研究分担者という形では協力者を揃えていない。但し、研究協力者としては、ポスドクの喬宇氏、また、本学博士学生の斎藤大輔氏、修士学生の鈴木雅之氏に依頼している。また、言語聴覚士（所属は清瀬市福祉センター）の桜庭京子氏とは長年に渡る協力体制を築いており、今後とも構造表象と言語障害、発達障害の関連性について種々の議論を行なうことになっている。



## これまでに受けた研究費とその成果等

本欄には、研究代表者がこれまでに受けた研究費（科学研究費補助金、所属研究機関より措置された研究費、府省・地方公共団体・研究助成法人・民間企業等からの研究費等。なお、現在受けている研究費も含む。）による研究成果等のうち、本研究の立案に生かされているものを選定し、科学研究費補助金とそれ以外の研究費に分けて、次の点に留意し記述してください。

- ① それぞれの研究費毎に、研究種目名（科学研究費補助金以外の研究費については資金制度名）、期間(年度)、研究課題名、研究代表者又は研究分担者の別、研究経費（直接経費）を記入の上、研究成果及び中間・事後評価（当該研究費の配分機関が行うものに限る。）結果を簡潔に記述してください。
- ② 科学研究費補助金とそれ以外の研究費は線を引いて区別して記述してください。

- **特定領域研究「情報学」平成13年～平成17年（公募研究）「人間による音声情報処理過程の分析とそれを応用した音声対話インターフェイスの構築」（研究代表者，27,200千円）**

人間は五感を用いながら、様々な情報を環境から得ている。ここでは聴覚に絞り、また対象を音声に絞った上で、非言語的情報（話者、年齢、性別等）や、パラ言語的情報（感情、意図等）の抽出過程を計算機上に実装した。知覚的年齢の推定に際しては、大規模音声データベースに対して聴取実験より年齢ラベリングを施し、それに基づいて各年代の音響モデルを作成し、知覚的年齢推定を実装した。また、パラ言語情報の推定に対しては、音声アフォーダンスの基盤ともなる音声の構造的表象を用いて分析を行ない、より高精度な話者意図の推定を実現した。音声に不可避的に混入する非言語的な特徴（話者、年齢、性別、マイク特性など）を一切含まない音声の物理表象への着眼は、本公募研究（情報学）においてその発想を得た。上記のパラ言語情報抽出に対する基礎的検討はその一例である。

- **特定領域研究「情報爆発」平成18年（公募研究）「構造不変の定理に基づく音声アフォーダンスの提案とそれに立脚した音声認識系の構築」（研究代表者，3,300千円）**

申請者が「構造不変の定理」と呼ぶ数学定理が導出する音声の構造的表象（音声アフォーダンス）による音声認識系の構築に着手し始めた。理論的進展としては、本表象が非線形変換を含む幅広い（ほぼ全ての）変換群に対して不変なる構造的情報表象を与えることが証明されたこと、幼児の言語獲得（音声模倣）を説明する一つの数理モデルとして妥当であること、自閉症、失読症が示す一部の症例の説明モデルとして妥当であること、を示せたことなどが挙げられる。工学的進展としては、孤立母音系列発声のみならず、連続発声された音声の構造化についても一定の成果が確認されたこと、などがある。

- **特定領域研究「情報爆発」平成19年～平成20年（公募研究）「構造不変の定理に基づく音声アフォーダンスの提案とそれに立脚した音声認識系の構築」（研究代表者，7,500千円）**

18年度の成果を踏まえ、連続発声された母音系列単語、更には、音素バランス単語に対する孤立単語音声認識をタスクとして、必要な要素技術を構築した。構造不変の定理は、如何なる写像に対しても不変性を保証するため、不変性が強すぎるという問題を持つ。この不変性を適切に制約する必要がある、これを次元分割を行なうことで実装した。その結果として、母音系列単語に対しては、話者の違いを完全に消した形で音声認識を実装することが可能となり、音素バランス単語についても、無ミスマッチ時での精度劣化を抑えつつ、ミスマッチが大きい時の精度を飛躍的に向上させることができた。この3年間の研究成果をより発展させ、構造表象に基づく音声認識系を完成させるために、改めて、研究助成金の申請を行なうものである。

領域略称名	情報爆発 IT 基盤	研究機関名	東京大学	研究代表者氏名	峯松 信明
-------	------------	-------	------	---------	-------

## 人権の保護及び法令等の遵守への対応（公募要領 6 頁参照）

本欄には、研究計画を遂行するにあたって、相手方の同意・協力を必要とする研究、個人情報の取り扱いの配慮を必要とする研究、生命倫理・安全対策に対する取組を必要とする研究など法令等に基づく手続きが必要な研究が含まれている場合に、どのような対策と措置を講じるのか記述してください。

例えば、個人情報を伴うアンケート調査・インタビュー調査、患者から提供を受けた試料の使用、ヒト遺伝子解析研究、組換え DNA 実験、動物実験など、研究機関内外の倫理委員会等における承認手続きが必要となる調査・研究・実験などが対象となります。

なお、該当しない場合には、その旨記述してください。

該当しない。

## 研究経費の妥当性・必要性

本欄には、「研究計画・方法」欄で述べた研究規模、研究体制等を踏まえ、様式「特定（公募）－ 9」以降に記入する研究経費のうちで、特に重要な経費の妥当性・必要性・積算根拠について記述してください。

また、研究計画のいずれかの年度において、各費目（設備備品費、旅費、謝金等）が全体の研究経費の 90% を超える場合及びその他の費目で、特に大きな割合を占める経費がある場合には、当該経費の必要性（内訳等）を記述してください。

平成 21 年度の研究計画は、1) 多様性を表現する写像関数群の推定方法の確立、2) 共鳴音と非共鳴音を分割した上での構造化の検討、3) 構造表象を通して考える韻律的特徴のパターン表象、4) 各種認知障害と構造表象との関連に関する調査・検討、である。購入予定物品は基本的に、これらの研究項目を推進するために不可欠な物品である。特に 1) の検討に際しては当初考えていたよりも遥かに高い計算機性能が必要であることが分かって来た。これは、構造表象の応用先を音声に留めず、視覚や身体運動にまで拡張しながら検討を進めていることも一つの理由であるが、音声に限った場合でも従来想定していた以上の計算機能力が必要となる。話者の多様性を必要十分に記述するための一般的な枠組みを検討するため、膨大な話者数の音声データを HDD 上に搭載し、各話者空間を多様体として実装した上で多様体間の写像を推定する必要がある。

2) や 3) に関しては、構造表象に基づく音響照合を各種検討していくことになる。そのための作業を円滑に行なうために、デスクトップ PC が必要となる。音声分析用 PC がこれに相当する。

4) 自閉症・失読症者の情報処理モデルの基礎検討であるが、これは、情報収集にかかる費用を計上している。本年度、情報収集を目的として国内外に出張し（国外としては、アメリカを 2 度訪問している。非音声工学者との議論が目的であり、脳科学、言語障害、言語学、進化人類学など様々な分野の研究者との議論を繰り返している）、また、情報源として種々の図書を購入し、提案手法の、工学のみならず科学的な妥当性の議論を繰り返している。図書購入費、及び、出張費（国内外）として計上しているが、恐らくこれだけでは不十分であると思われる（本年度、既に図書費だけで 50 万円を超えている）。情報収集費に関しては、他の助成金なども用いて補填していく予定である。

21 年度にデータグローブの購入を予定している。これは、22 年度に計画している身体運動と音運動間の、メディア普遍表象の初期検討を行なうためである。旧式のデータグローブによる検討は既に開始しており、研究の進捗状況を考慮して、時間・空間分解能の高いデータグローブを購入する。

設備備品費の明細						
(金額単位：千円) <span style="float: right; border-left: 1px solid black; padding-left: 5px;">                         多数の図書、資料を購入する場合は「西洋中世政治史関係図書」のようにある程度、図書、資料の内容が判明するよう                          な表現で記入してください。また、機械器具の場合は、単に〇〇〇一式とするだけでなくその内訳も記入してください。                     </span>						
年度	品名・仕様	数量	単価	金額	主として使用する研究者及び設置機関名	購入予定時期
21	音声処理装置・DELL PowerEdge1950	1	400	400	東京大	9月
	音声分析用PC・DELL XPS 720	1	300	300	東京大	11月
	5DT社 Data Glove	1	600	600	東京大	9月
	好感度マイクロフォン	10	30	300	東京大	10月
	デジタル音声収録機・マランツ MD671	4	150	600	東京大	10月
	計				2,200	
22	RAID ディスク・Apple X-RAID	1	900	900	東京大	8月
	評価実験用PC・DELL XPS 720	2	300	600	東京大	10月
	計			1,500		

**消耗品費等の明細** (記入に当たっては、特定領域研究研究計画調書作成・記入要領（公募研究（新規）用）を参照してください。)  
 (金額単位：千円)

年度	消耗品費		旅 費		謝 金 等		そ の 他	
	品 名	金 額	事 項	金 額	事 項	金 額	事 項	金 額
平成 21 年度	言語障害・認知障害関係図書	200	(国内) 研究打合	100	音声収録謝金 (20人×10時間×1千円) 情報提供謝金 (2人×10時間×5千円)	200	会場費 論文別刷り代	100 200
	認知科学・脳科学関係図書	200	研究発表	300				
	CD/DVD メディア	100	(外国) 研究調査	300				
	内蔵 HD	100	研究発表	500				
	増設メモリ	100						
	計	700	計	1,200	計	300		
平成 22 年度	言語障害・認知障害関係図書	200	(国内) 研究打合	100	音声収録謝金 (20人×10時間×1千円) 情報提供謝金 (2人×10時間×5千円)	200	脳計測/聴覚計測機器借用料 会場費 論文別刷り代	600 100 200
	認知科学・脳科学関係図書	200	研究発表	300				
	CD/DVD メディア	100	(外国) 研究調査	300				
	内蔵 HD	100	研究発表	500				
	増設メモリ	100						
	計	700	計	1,200	計	300		
領域略称名	情報爆発 IT 基盤		研究機関名	東京大学		研究代表者氏名	峯松 信明	

# 研究業績

特定(公募) - 10 - (1)

本欄には、研究代表者及び研究分担者が最近5カ年間に発表した論文、著書、産業財産権等、招待講演のうち、本研究に関連する重要なものを選定し、現在から順に発表年次を過去にさかのぼり、発表年(暦年)毎に線を引いて区別(線は移動可)し、通し番号を付して3頁以内で記入してください。なお、学術誌へ投稿中の論文を記入する場合は、掲載が決定しているものに限ります。

また、必要に応じて、連携研究者の研究業績についても記入することができます。記入する場合には、二重線を引いて区別(二重線は移動可)し、研究者毎に、現在から順に発表年次を過去にさかのぼり記入してください。(発表年毎に線を引く必要はありません。)

発表年	研究代表者・ 分担者氏名 (所属研究機関・ 部局・職)	発表論文名・著書名等 (例えば発表論文の場合、論文名、著者名、掲載誌名、査読の有無、巻、最初と最後の頁、発表年(西暦)について記入してください。) (以上の各項目が記載されていれば、項目の順序を入れ替えても可。著者名が多数にわたる場合は、 <u>主な著者を数名記入し以下を省略(省略する場合、その員数と、掲載されている順番を○番目と記入)しても可。なお、研究代表者には二重下線、研究分担者には一重下線、連携研究者には点線の下線を付してください。また、corresponding author には左に*印を付してください。)</u>
2008	峯松信明(東京大学・工学系研究科・准教授)	<p>1. *<u>峯松信明</u>, “音声言語運用が要求する認知的能力と音声言語工学が構築した計算論的能力”, 音声言語シンポジウム 10 周年記念特別招待講演 (2008, 招待講演)</p> <p>2. *<u>N. Minematsu</u>, T. Nishimura, “Consideration of infants’ vocal imitation through modeling speech as timbre-based melody,” in New Frontiers in Artificial Intelligence, LNAI4914, 26–39, Springer (2008, 査読有)</p> <p>3. 村上隆夫, *<u>峯松信明</u>, 廣瀬啓吉 “音声の構造的表象に基づく日本語孤立母音系列に対する音声認識”, 電子情報通信学会論文誌, J91-A, 2, 181–191 (2008, 査読有)</p> <p>4. *<u>峯松信明</u>, “「あ」という声を聞いて母音「あ」と同定する能力は音声言語運用に必要か?”, 日本語学, 4, 187–197, 明治書院 (2008, 査読有)</p> <p>5. *<u>N. Minematsu</u>, “Speech as music –Yet another framework for speech recognition,” ESF-JSPS Frontier Science Conference Series for Young Researchers – Experimental Cognitive Robotics – (2008, 招待講演)</p> <p>6. *<u>N. Minematsu</u>, “Training of pronunciation as learning of the sound system embedded in the target language,” Proc. The 8th Phonetic Conference of China and Int. Symposium on Phonetic Frontiers (2008, 招待講演)</p> <p>7. D. Saito, S. Asakawa, *<u>N. Minematsu</u>, and K. Hirose, “Structure to speech – speech generation based on infant-like vocal imitation –,” Proc. INTERSPEECH, 1837–1840 (2008, 査読有)</p> <p>8. D. Saito, *<u>N. Minematsu</u>, and K. Hirose, “Decomposition of rotational distortion caused by VTL difference using eigenvalues of its transformation matrix,” Proc. INTERSPEECH, 1361–1364 (2008, 査読有)</p> <p>9. Y. Qiao and *<u>N. Minematsu</u>, “f-divergence is a generalized invariant measure between distributions,” Proc. INTERSPEECH, 1349–1352 (2008, 査読有)</p> <p>10. *<u>N. Minematsu</u> and T. Nishimura, “Speech as timbre-based melody – What in parents’ voices do infants imitate acoustically? –,” Proc. Int. Conf. Language, Music, and the Mind, Paper Session (2008, 査読有)</p> <p>11. *<u>N. Minematsu</u>, T. Nishimura, D. Saito, S. Asakawa, Y. Qiao, “Holistic and prosodic representation of the segmental aspect of speech,” Proc. Int. Conf. Speech Prosody, 169–172 (2008, 査読有)</p> <p>12. D. Saito, R. Matsuura, S. Asakawa, *<u>N. Minematsu</u>, K. Hirose, “Directional dependency of cepstrum on VTL,” Proc. ICASSP, 4485–4488 (2008, 査読有)</p> <p>13. S. Asakawa, *<u>N. Minematsu</u>, K. Hirose, “Multi-stream parameterization for structural speech recognition,” Proc. ICASSP, 4097–4100 (2008, 査読有)</p> <p>14. 國越晶, 喬宇, *<u>峯松信明</u>, 廣瀬啓吉, “空間写像に基づく手の動きを入力とした音声生成系”, 日本音響学会講演集, 1-Q-23, 375–376 (2008, 査読無)</p> <p>15. 喬宇, *<u>峯松信明</u>, “変換不変性を有するダイバージェンスとその一般形”, 信学技報, SP2008-51, 49–54 (2008, 査読無)</p> <p>16. 齋藤大輔, 朝川智, *<u>峯松信明</u>, 廣瀬啓吉, “構造的表象からの音声合成とそれに基づく音声模倣に関する検討” 信学技報, SP2008-40, 115–120 (2008, 査読無)</p> <p>17. 鈴木雅之, 朝川智, 喬宇, *<u>峯松信明</u>, 廣瀬啓吉, “スペクトル特徴量を用いた音声の構造的表象に関する実験的検討”, 信学技報, SP2008-32, 73–78 (2008, 査読無)</p>

# 研究業績

特定(公募) - 10 - (2)

本欄には、研究代表者及び研究分担者が最近5か年間に発表した論文、著書、産業財産権等、招待講演のうち、本研究に関連する重要なものを選定し、現在から順に発表年次を過去にさかのぼり、発表年(暦年)毎に線を引いて区別(線は移動可)し、通し番号を付して3頁以内で記入してください。なお、学術誌へ投稿中の論文を記入する場合は、掲載が決定しているものに限ります。

また、必要に応じて、連携研究者の研究業績についても記入することができます。記入する場合には、二重線を引いて区別(二重線は移動可)し、研究者毎に、現在から順に発表年次を過去にさかのぼり記入してください。(発表年毎に線を引く必要はありません。)

発表年	研究代表者・ 分担者氏名 (所属研究機関・ 部局・職)	発表論文名・著書名等 (例えば発表論文の場合、論文名、著者名、掲載誌名、査読の有無、巻、最初と最後の頁、発表年(西暦)について記入してください。) (以上の各項目が記載されていれば、項目の順序を入れ替えても可。著者名が多数にわたる場合は、 <u>主な著者を数名記入し以下を省略(省略する場合、その員数と、掲載されている順番を○番目と記入)しても可。なお、研究代表者には二重下線、研究分担者には一重下線、連携研究者には点線の下線を付してください。また、corresponding author には左に*印を付してください。)</u>
		18. 朝川智, 喬宇, <u>*峯松信明</u> , 広瀬啓吉, “判別分析と構造構造表象を用いた話者の多様性に超頑健な音声認識”, 日本音響学会講演集, 2-P-3, 113-116 (2008, 査読無)
2007	峯松信明(東京大学・工学系研究科・准教授)	19. <u>*峯松信明</u> , “小特集「言語障害を通して再考する音声言語情報処理」にあたって”, 日本音響学会誌, 63, 7, 363-364 (2007, 査読有) 20. 朝川智, <u>*峯松信明</u> , 広瀬啓吉, “音声の構造的表象に基づく英語学習者発音の音響的分析”, 電子情報通信学会論文誌, J90-D, 5, 1249-1262 (2007, 査読有) 21. Y. Qiao, S. Asakawa, <u>*N. Minematsu</u> , “Random discriminant structure analysis for continous Japanese vowel recognition,” Proc. Int. Workshop on Automatic Speech Recognition and Understanding, 576-581 (2007, 査読有) 22. <u>*N. Minematsu</u> , K. Kamata, S. Asakawa, T. Makino, and K. Hirose, “Structural representation of the pronunciation and its use for classifying Japanese learners of English,” Proc. ISCA Tutorial and Research Workshop on Speech and Language Technology in Education, CD-ROM (2007, 査読有) 23. <u>*N. Minematsu</u> , “Are learners myna birds to the averaged distributions of native speakers? – a note of warning from a serious speech engineer –,” Proc. ISCA Tutorial and Research Workshop on Speech and Language Technology in Education, CD-ROM (2007, 査読有) 24. <u>*N. Minematsu</u> , K. Kamata, S. Asakawa, T. Makino, T. Nishimura, and K. Hirose, “Structural assessment of language learners’ pronunciation,” Proc. INTERSPEECH, 210-213 (2007, 査読有) 25. S. Asakawa, <u>*N. Minematsu</u> , and K. Hirose, “Automatic recognition of connected vowels only using speaker-invariant representation of speech dynamics,” Proc. INTERSPEECH, 890-893 (2007, 査読有) 26. <u>*N. Minematsu</u> , K. Sakuraba, and T. Nishimura, “Development of a spoken word recognizer without phonemic awareness – Is this machine a Dyslexia simulator? –,” Proc. The 2nd Riken and Oxford Joint International Symposium – Reading and Dyslexia in Different Languages –, 49-51 (2007, 査読有) 27. <u>*N. Minematsu</u> , “From fragile substances to robust contrasts – Yet another framework for speech recognition –,” Seminar talk at Microsoft Research Asia and Institute of Linguistics, Chinese Academy of Social Sciences (2007, 招待講演) 28. <u>*N. Minematsu</u> , “Are learners myna birds to the averaged distribution of native speaker? – A new framework for CALL –,” Seminar talk at University of Science and Technology of China (2007, 招待講演) 29. <u>*N. Minematsu</u> , “Speech as multidimensional music,” Seminar talk at University College of London (2007, 招待講演) 30. <u>*峯松信明</u> , “音楽としての音声～構造的不変性に基づく人間の音認知～”, 視聴覚情報研究会 (AVIRG) (2007, 招待講演) 31. 齋藤大輔, 松浦良, 朝川智, <u>*峯松信明</u> , 広瀬啓吉, “ケプストラムの声道長依存性に関する幾何学的考察”, 信学技報, SP2007-128, 189-194 (2007, 査読無) 32. Y. Qiao, S. Asakawa, and <u>*N. Minematsu</u> , “Recognition of connected Japanese vowel utterances using random discriminant structure analysis,” IEICE Technical Report, SP2007-89, 19-24 (2007, 査読無)

本欄には、研究代表者及び研究分担者が最近5か年間に発表した論文、著書、産業財産権等、招待講演のうち、本研究に関連する重要なものを選定し、現在から順に発表年次を過去にさかのぼり、発表年(暦年)毎に線を引いて区別(線は移動可)し、通し番号を付して3頁以内で記入してください。なお、学術誌へ投稿中の論文を記入する場合は、掲載が決定しているものに限ります。

また、必要に応じて、連携研究者の研究業績についても記入することができます。記入する場合には、二重線を引いて区別(二重線は移動可)し、研究者毎に、現在から順に発表年次を過去にさかのぼり記入してください。(発表年毎に線を引く必要はありません。)

発表年	研究代表者・ 分担者氏名 (所属研究機関・ 部局・職)	発表論文名・著書名等 (例えば発表論文の場合、論文名、著者名、掲載誌名、査読の有無、巻、最初と最後の頁、発表年(西暦)について記入してください。) (以上の各項目が記載されていれば、項目の順序を入れ替えても可。著者名が多数にわたる場合は、 <u>主な著者を数名記入し以下を省略(省略する場合、その員数と、掲載されている順番を○番目と記入)しても可。なお、研究代表者には二重下線、研究分担者には一重下線、連携研究者には点線の下線を付してください。また、corresponding author には左に*印を付してください。)</u>
		<p>33. 齋藤大輔, 朝川智, <u>*峯松信明</u>, 広瀬啓吉, “構造的表象からの音声生成に関する基礎的検討”, 信学技報, SP2007-80, 55-60 (2007, 査読無)</p> <p>34. 鎌田圭, 朝川智, <u>*峯松信明</u>, 牧野武彦, 広瀬啓吉, “学習者による米語母音発音に対する絶対的評価と相対的評価間の相関分析”, 信学技報, SP2007-36, 73-78 (2007, 査読無)</p> <p>35. <u>*峯松信明</u>, 西村多寿子, 櫻庭京子, 朝川智, 齋藤大輔, “孤立音[あ]を聞いて/あ/と同定する能力は音声言語に必要か?”, 信学技報, SP2007-30, 37-42 (2007, 査読無)</p> <p>36. <u>*峯松信明</u>, 西村多寿子, 朝川智, 櫻庭京子, 齋藤大輔, “要素論から全体論へ～全体から入る音声情報処理への招待～”, 情報処理学会音声言語情報処理研究会, 2007-SLP-67-14, 75-80 (2007, 査読無)</p> <p>37. <u>*峯松信明</u>, 西村多寿子, 櫻庭京子, 朝川智, “孤立音を聞いて音韻同定できる能力は音声言語運用に必要か?”, 日本音声学会全国大会予稿集, 135-140 (2007, 査読有)</p> <p>38. <u>*峯松信明</u>, 西村多寿子, “多次元音楽としての音声モデリングと音声模倣”, 人工知能学会全国大会講演論文集, 1F2-3, 1-44 (2007, 査読無)</p> <p>39. <u>*峯松信明</u>, 朝川智, 広瀬啓吉, “線形・非線形変換不変の構造的情報表象とそれに基づく音声の音響モデリングに関する理論的考察”, 日本音響学会講演集, 1-P-12, 147-148 (2007, 査読無)</p> <p>40. <u>*峯松信明</u>, 櫻庭京子, 西村多寿子, “非言語的要因に不変な音声の構造的表象に基づく失読症・自閉症に関する一考察”, 日本音響学会講演集, 2-8-15, 365-366 (2007, 査読無)</p>
2006	峯松信明(東京大学・工学系研究科・准教授)	<p>41. <u>*峯松信明</u>, “音楽としての音声 ～構造的不変性に基づく人間の音認知～”, IBM 東京基礎研究所講演会 (2006, 招待講演)</p> <p>42. <u>*峯松信明</u>, “音声の構造的表象を通して考察する幼児の音声模倣～幼児は親の声の何を真似ようとしているのか～”, 人工知能学会幼児のコモンセンス知識研究会講演会 (2006, 招待講演)</p> <p>43. <u>*峯松信明</u>, “音声の構造解析とゲノム～音声はゲノムか, ゲノムは音声か～”, 21世紀 COE プログラム「言語から読み解くゲノム」特別講義 (2006, 招待講演)</p> <p>44. <u>*峯松信明</u>, “脳の情報処理モデルと音声の認識・生成”, 電子情報通信学会ヒューマンコミュニケーショングループシンポジウム (2006, 招待講演)</p> <p>45. <u>*N. Minematsu</u>, “Structural representation of the pronunciation and its application to computer-aided language learning,” ASA &amp; ASJ Joint meeting (2006, 招待講演)</p> <p>46. <u>*N. Minematsu</u>, S. Asakawa, and K. Hirose, “Structural representation of the pronunciation and its use for CALL,” Int. Workshop on Spoken Language Technology, pp.126-129 (2006)</p> <p>47. <u>*N. Minematsu</u>, T. Nishimura, K. Nishinari, and K. Sakuraba, “Theorem of the invariant structure and its derivation of speech Gestalt,” Proc. Int. Workshop on Speech Recognition and Intrinsic Variations, pp.47-52 (2006)</p>
	連携研究者氏名 (所属研究機関・部局・職)	発表論文名・著者名 等 (研究代表者及び研究分担者の研究業績として上欄に記載したものは記載しないでください。)

**研究費の応募・受入等の状況・エフォート**

本欄は「研究資金の不合理な重複や過度の集中にならず、研究課題が十分に遂行し得るかどうか」を判断する際に参照するところであり、研究代表者の、応募時点における、（１）応募中の研究費、（２）受入予定の研究費、（３）その他の活動、について、次の点に留意し記入してください。なお、複数の研究費を記入する場合は、線を引いて区別して記入してください。

- ① 「エフォート」欄には、年間の全仕事時間を100%とした場合、そのうち当該研究の実施等に必要となる時間の配分率（%）を記入してください。
- ② 「応募中の研究費」欄の先頭には、本応募研究課題を記入してください。
- ③ 科学研究費補助金の「特定領域研究」にあつては、「計画研究」、「公募研究」の別を記入してください。
- ④ 所属研究機関内で競争的に配分される研究費についても記入してください。

**（１） 応募中の研究費**

資金制度・研究費名（研究期間・配分機関等名）	研究課題名 （研究代表者氏名）	役割（代表・分担の別）	平成21年度研究経費（期間全体の額）（千円）	エフォート（%）	研究内容の相違点及び他の研究費に加えて本応募研究課題に応募する理由
【本応募研究課題】 特定領域研究 公募研究 (H21～H22)	構造不変の定理に基づく 音声アフォーダンスの提案とそれに基づく音声認識系の構築	代表	4,700 (9,300)	20	/
挑戦的萌芽研究 (H21～H22) 東京大学	幼児の音声模倣と音声ゲシュタルトに着眼した構造表象に基づく音声生成系の構築	代表	2,650 (5,000)	—	音声合成（生成）に主眼を置いた構造表象の応用研究であり、音声を受容するための技術構築とは異なる。
若手 (S) (一般) (H21～H25) 東京大学	構造不変性に基づくメディア普遍の運動表象とそれに基づく脳型メディア情報処理	代表	17,620 (99,900)	35	構造表象を、身体・視覚・聴覚運動を統合した動き表象として考え、処理系を構築する。購入予定物品が申請課題と大きく異なる。
基盤研究 (B) (H21～H23) 東京大学	生成過程モデルに基づく表現力豊かな多言語音声合成とそれによる音声自動翻訳（広瀬啓吉）	分担	600 (1,800)	5	複数言語の音声合成・音声翻訳に主眼を置いた研究であり、本申請研究との関係は無い。
基盤研究 (B) (H21～H23) 東京国際大学	音読とシャドーイングに音声情報処理技術を活用した外国語能力評価システムの開発（山内豊）	分担	3,000 (6,000)	5	シャドーイングという外国語学習に特化した発音・聴取訓練方法の自動化に関する研究であり、本申請課題との関係は無い。

研究費の応募・受入等の状況・エフォート（つづき）					
（２）受入予定の研究費					
資金制度・研究費名・研究期間（配分機関等名）	研究課題名（研究代表者氏名）	役割（代表・分担の別）	平成21年度研究経費（期間全体の額）（千円）	エフォート（％）	研究内容の相違点及び他の研究費に加えて本応募研究課題に応募する理由
基盤研究 (B) (H20 ～ H22) 東京大学	音声の分節的・韻律的特徴を包含する発音の構造的表象に基づく外国語教育・学習支援	代表	4,800 (14,700)	—	構造表象に基づく外国語発音学習支援環境の構築であり、申請課題とは目的が大きく異なる。
特定領域研究 (計画班) (H18 ～ H22) 東京大学	多様な目的に適した形態素解析システム用電子化辞書の開発（伝康晴）	分担	2,500 (13,300)	10	音声合成のテキストモジュール開発を目的とした研究であり、申請課題と直接的な関係は無い。
<b>（３）その他の活動</b> 〔 上記の応募中及び受入予定の研究費による研究活動以外の職務として行う研究活動や教育活動等のエフォートを記入してください。 〕				25	/
<b>合 計</b> （上記（１），（２），（３）のエフォートの合計）				100 (%)	