

機関番号	研究種目番号	審査区分番号	細目番号	分割番号	整理番号
12601	11	-	1006	A	0003

## 平成21年度 (2009年度) 挑戦的萌芽研究 研究計画調書

平成 20 年 11 月 4 日  
2 版

### 新規

研究種目	挑戦的萌芽研究						
分野	総合領域						
分科	情報学						
細目	知覚情報処理・知能ロボティクス						
細目表 キーワード	音声情報処理						
細目表以外の キーワード	発達心理学・進化人類学						
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	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						
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	2,650	1,200	400	850	200	0
	平成22年度	2,350	750	400	850	200	150
	平成23年度	0	0	0	0	0	0
	総計	5,000	1,950	800	1,700	400	150
開示希望の有無	審査結果の開示を希望する						



## 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、冒頭にその要旨を記述した上で、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（基盤研究等）における審査及び評価に関する規程」（公募要領 52～99 頁参照）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進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 当該分野における本研究の学術的な特色及び予想される結果と意義

発達心理学では「幼児の言語獲得は、親の声の音声模倣を通して行なわれる」と言われる。しかし、親の声の「物真似」をする幼児はいない（ちなみに、九官鳥は飼い主の個人性まで「物真似」する）。つまり、彼らは声そのものは一切真似ない。声の何を真似るのだろうか？幼児の音韻意識は未熟であるため（例えば、日本語に5個の母音があることを知らない）、親の声を個々のモーラ（平仮名）として同定し、個々のモーラを声に変換する、という解釈は不適切である。発達心理学は「個々のモーラを獲得する前に、語全体の音形、音声ゲシュタルトを獲得する」と主張する。となれば、その音声ゲシュタルトには話者の情報は含まれていないはずである。含まれていれば、幼児は、父親そっくりの「声」を出そうと努力するはずである。申請者は、この「音声ゲシュタルト」の音響的定義を数多くの発達心理学研究者に対して調査したが、その定義を語る者はいなかった [1]。

**本研究の目的は、話者不変な音声表象である音声ゲシュタルトを物理的に実装し、それに基づく音声模倣（獲得）を行なう能力を計算機の上に実装することで、音声生成系を構築することを目的とする。**

動物学にまで目を向けると、音声模倣や音声学習（親の声に対して、子供がそれを積極的に真似る行為）は非常に限られた種でしか観測されない。ヒト以外での霊長類では観測されない。観測される種は、例えば小鳥やクジラなどがあるが、彼らの音声模倣は「声真似」である。ヒトの音声模倣は個体サイズ（体格）を超えた模倣をしていく。これは、ヒト固有の音声模倣形態であり、動物学的には非常に不思議な模倣形態である [2]。

ヒトや動物の音声模倣を念頭に置いて従来の音声合成装置を鑑みるに、それらは、提供されたある話者の音声に対して、その話者性までも見事に再現する。即ち、従来の音声合成装置は動物（九官鳥）や物真似タレントの模擬装置でしかない。本研究は、その動作が「人間のように見える」機械を作るのではなく、「ヒトしか持ち合わせていない模倣能力を計算機に宿すことで構築される音声生成系の試作」が目的である。

① 音声合成研究は非常に長い歴史を持つが、申請者は、九官鳥模擬装置以外の実装系を知らない。ある男性の「野太い声」を学習データとして与えたにも拘らず、どうしても「可愛い声」しか出ない音声合成装置を申請者は知らない。本研究で目指すのは、このような音声生成系である。話者不変の「音声ゲシュタルト」の物理的実装であるが、既に申請者が検討しており、音声認識研究において、優れた成果を出している [3][4]。個々の音事象の「音そのもの」をモデル化するのではなく、「音と音の差異（コントラスト）」を単語全体に渡って全て計測すれば、それは距離行列となり（＝一つの幾何学的構造を規定する）、コントラスト計測に f-divergence を用いれば、完全な変換（写像）不変量であるため、結局、得られる距離行列（音声構造）は話者不変となる。この音声の全体（ゲシュタルト）表象を用いて孤立単語認識を実装したところ、**四千人以上の話者の声を用いて構築された音の実体モデル（従来法）では、認識率が 1.4%となる条件下（大人の声を用いてモデルを作り、小人の声を認識させる）においてでも、数名の話者から構築した全体的な構造モデルでは、90%を超えた。**

既にこの音声の構造表象（音声ゲシュタルト）を用いた音声生成系（音声模倣系）の予備的検討は完了しており [5]、本研究では、この音声生成系をより洗練させ、最終的には、発達心理学研究者の評価に耐えうるシステムを構築することが目的である。幼児の音声模倣をどこまで再現できるのか、を本研究では追求してみたい。

② 文献 [5] では、孤立母音の系列や、母音列を連続発声した発話を対象として、父親の太い声から語ゲシュタルトを抽出し、それに対して、短い声道しか持ち合わせていない幼児が、その語ゲシュタルトを自らの小さな口で再生するプロセスを（音響的に）計算機に実装した。その結果、どのような発声に対しても、「可愛い」声でしか模倣できないシステムが構築された。本研究では、子音や韻律的特徴など、様々な未解決問題について検討し、最終的には、幼児の音声模倣の再現性について、発達心理学研究者の評価に耐えうるシステムを構築する。

③ 発達心理学的には、幼児の音声模倣プロセスに対する物理モデルを提供できる。さらに、申請者が提唱している音声ゲシュタルトの物理的解釈の妥当性を、従来とは異なる観点から検証できる。音声工学的には、従来試みられてことが無い全く新しい音声生成系の構築パラダイムを提供できる。なお、文献 [1] は高く評価され、論文賞を獲得している。様々な分野においてその成果が待たれている研究プロジェクトであると自負している。

[1] N. Minematsu *et al.*, "A consideration of infants' vocal imitation," LNAI4914, 26-39, Springer, 2008

[2] 岡ノ谷一夫, "小鳥の歌と言語：共通する進化メカニズム", 音講論, 1-7-15, 1555-1556, 2008

[3] S. Asakawa *et al.*, "Multi-stream parameterization for structural speech recognition," ICASSP, 4097-4100, 2008

[4] 峯松信明, "音声言語運用が要求する認知的能力と音声言語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

[5] D. Saito *et al.*, "Speech generation based on infant-like vocal imitation," INTERSPEECH, 1837-1840, 2008

研究の斬新性・チャレンジ性

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 本研究が、どのような点で斬新なアイデアやチャレンジ性を有しているか
- ② 本研究が、新しい原理の発展や斬新な着想や方法論の提案を行うものである点、または成功した場合に卓越した成果が期待できるものである点等

①&② 本研究プロジェクトの斬新性・チャレンジ性は極めて高い。

まず、音声の全体的・構造的表象（音声ゲシュタルト）であるが、これは音声科学・音声工学の古典的かつ未解決の問題である「音声の物理的多様性と音声知覚の不変性」を解決する方法として近年提案されている非常に新しい表象（音声の捉え方）である。話者の体格の違いは、声道形状の個人性を生じ、その結果、音声（音色）の音響特性は人によって大きく異なる。しかし、これらの違いは静的なバイアス項でしかない。色知覚の不変性（恒常性）や、メロディー知覚の不変性はいずれも、刺激と刺激のコントラストを捉えることで静的なバイアス項を相殺していると説明される。個々の刺激がどのような絶対的物理特性を有しているのか、ではなく、各刺激が他刺激群とどのような関係性を保持しているのか、によって各刺激の知覚イメージが形成される。興味深いのは、色知覚の不変性は昆虫でも観測される進化的に古い能力であるが、メロディー知覚の不変性（音高の相対音感）はマカク猿でも観測されない極めて新しい能力である、という事実である。音声ゲシュタルトは、これら生態心理学、進化心理学的な知見を背景に提案された新しい音声表象（音色の相対音感）である。

個々の刺激の絶対的特性ではなく「個々の刺激と他刺激群との間に成立している関係性が各刺激の知覚イメージを形成する」という解釈は、上記したように認知心理学や、更には認知言語学、古くは構造言語学において見られる議論と一致する。しかし、この考え方を、数学的・物理学的に定義し、工学的な実用アプリケーション開発を可能とする枠組みとして纏め上げたのは、申請者が世界で初めてである。

また、可逆かつ連続的な如何なる変換に対しても不変であるという f-divergence の特性は、位相幾何学 (Topology) における新たな位相不変量として考察できる。この新しく導かれた位相不変量に基づいて知覚の恒常性を説明する物理的な枠組みを構築している。このように本研究は理論数学的にも非常に興味深い不変量を導出しており、また、それを実世界の問題を解決するために応用している訳だが、異なるメディアの同種の問題にも応用可能であると考えている。即ち本プロジェクトは、**強固な、かつ、新しい数学的基盤**を持っており、本プロジェクトの成功は、この基盤の他分野（メディア）への応用可能性を一気に切り開く可能性がある。

「研究の目的」にも記した通り、多くの音声合成研究は（調音音声合成を覗き）、全て、学習話者から与えられた音声のある単位で切り刻み、素片の出現順序を変えて繋ぎ合わせて音声波形を生成する方法論に基づいている。その結果、学習話者の話者性までも見事に再現される。これは九官鳥的な音声模倣である。本研究は、この枠組みとは異なる枠組みの上に音声生成系を構築する試みであり、しかもそれは「ヒトのみが持つ個体サイズを超えた音声模倣能力の実装」を通して構築される系である。より具体的には、従来の音声合成研究は、与えられた音声資料の「音そのもの」を捉え、データベースとして保管し、適切な素片を選択して接続する方法論であった。本研究では、話者不変量となる「音と音の差異・コントラスト」を捉え、データベースとして保管し、適切な差異素片を選んで発話を構成する枠組みである。この時、差異素片群を組み上げても、結局それは話者不変の発声表象であり、具体的な声にはなっていない。声にするためには発話者の声道形状（デフォルト形状や長さ）が必要である。子供の声が可愛いのは、声道が短いからである。音声ゲシュタルトは、発声から、発声者の身体情報を取り除く。本研究で追求する枠組みは、音声ゲシュタルトに身体を戻すことで構造を具体的な音空間において復元する枠組みである。この時、身体が小さければ、可愛い声にしかならない。それだけである。このように、従来と全く異なる枠組みの上に音声生成系が実装できれば、従来の枠組みを相対的に検証することが可能となり、従来の音声合成技術のレベルアップにも間接的に寄与することが可能である。

更に、本研究の遂行は、音声模倣過程の物理的な情報処理モデルを提供することとなり、発達心理学研究に大きく貢献できる。更には、本研究の遂行は様々な発達障害研究にも応用可能であると考えている。例えば、重度の自閉症者の中には、九官鳥のように声真似をする自閉症者がいる。この場合、当然音声言語の運用は困難である。母親の声は理解できるが、父親の声は母親の声と異なるため、認識・理解できない自閉症者もいる。自閉症者の情報処理と、ヒト以外の霊長類の情報処理の類似性は動物学、発達障害学などで報告されているが、ヒト以外の霊長類が**音高**の相対音感を持っていない事実は、彼らや重度の自閉症者が**音色**の相対音感を持ち合わせていないと仮説を立てる十分な動機を与える。本研究の遂行によって、**音色**の相対音感による音声模倣プロセスの実装系（シミュレータ）が完成することになるが、この実装系を種々の条件下で動かすことによって「**音色**の相対音感モジュールが崩れると、音声模倣の様子がどう変わるのか」について一つの知見を与えることになる。これが、自閉症者の模倣形態と類似していれば、自閉症者の情報処理モデルを与えることに繋がる。

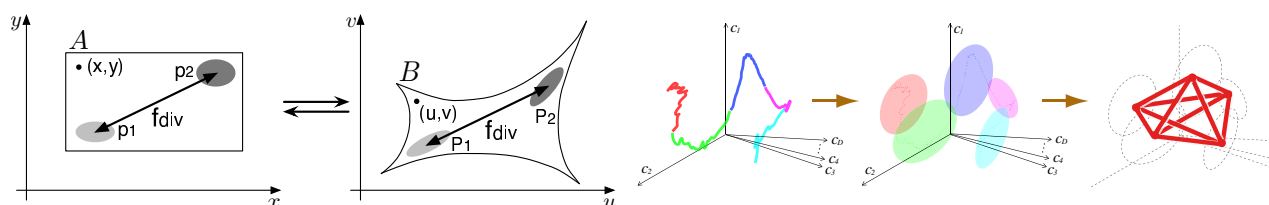
以上のように、本研究は、新しい数学的基盤に支えられ、進化心理学・生態心理学の比較的新しい知見に基づき、健常者の言語獲得プロセスのみならず、自閉症者の音情報処理モデルを与える可能性もあり、非常に多岐に渡る分野に対して新しい情報を発信することが可能となる。もちろん、音声工学的には、新しい音声合成系の構築パラダイムを提供することになる。本研究は、極めて高い、優れた斬新性・チャレンジ性を有している。

研究計画・方法

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、冒頭にその要旨を記述した上で、平成21年度の計画と平成22年度以降の計画に分けて、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。また、研究計画を遂行するための研究体制について、研究代表者及び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。さらに、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。

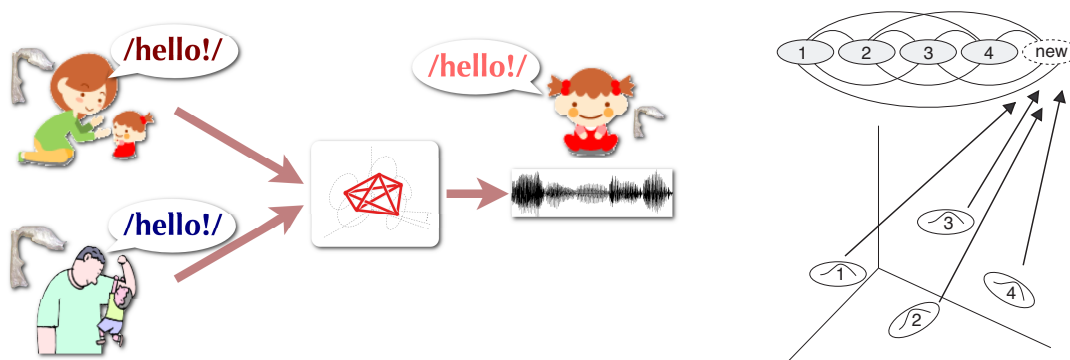
本申請課題は、研究代表者である峯松以外にも、研究室所属の博士学生1名、修士学生1名が参加予定である。また、連携研究者の梶川（認知発達科学）とも協力しつつ研究を進める予定である。

申請者が提唱している音声ゲシュタルトは、「構造的不変性」という数学的に保証された不変性を用いた音声表象である。ここで「構造的不変性」とは、観測対象に対して、如何なる連続的かつ可逆の線形・非線形写像（変換）を施しても、一切不変の量（情報）が、観測対象の「運動」の中にあることを数学的に証明した定理である。



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

個々の音事象の「音そのもの」をモデル化するのではなく、「音と音の差異（コントラスト）」を単語全体に渡って全て計測すれば、それは距離行列（＝一つの幾何学的構造を規定する）となり、コントラスト計測に f-divergence を用いれば、それは、完全な変換（写像）不変量となる。ある話者の音声を空間写像すれば、それは他話者の音声となるため、結局、得られる距離行列（音声構造）は話者不変となる。ある発声からの構造（語ゲシュタルト）抽出は、その話者の身体情報（声道のサイズや長さの情報）を音声から消失させることに等しい。逆に、構造に別話者の身体情報を戻すことで、構造は実際の声空間において定位されることとなり、結局、その身体を持った話者の声となる。構造に与えた話者の身体サイズが小さければ、それは幼児の音声模倣モデルとなる（下図参照）。



図：「構造表象を用いた幼児の音声模倣のモデル化」と「空間探索に基づく音声生成」

研究目的の欄に示した様に、既に、構造からの音声生成に関して初期検討を開始している。ここでは、実際に「小さな声道」という身体的な制約条件を課すのではなく、上図右に示す様に、既に幾つかの音が実声空間に定位されているとして、次の音を、既に定位された音からの距離（f-div. で計測）だけを用いて探索する問題として定式化した（図は、既に定位されている音1から音4と、次に来るべき音と各音*i*との距離の情報を元に、次の音を定位する様子を示している）。これは、例えば歌を歌う時に、歌い始めは音高制御が不安定だが、既に幾つかの音を出した後、スムーズに歌い続けることができる現象を、音色の空間で検討していることに相当する。本申請課題では、1) この枠組みの精緻化と、2) 幼児による実際の音声模倣時の音声資料の音響的・構造的な分析、3) 実際の身体情報（声道サイズ情報）を入力とした音声生成系の構築、4) 構築した音声生成系の発達心理学的評価、などを具体的な研究目的としている。以下、各テーマについて述べる。

平成21年度の研究計画・方法

1. 共鳴音・非共鳴音を含んだ音声に対する空間探索手法の構築【担当：峯松・博士】

構造からの音声生成における初期検討では、母音列のみからなる音声を対象としていた。共鳴子音（/m/, /n/, /w/, /r/, /l/, /y/など）は現状の枠組みでも生成可能であるが、非共鳴子音の場合（/s/, /t/, /ch/, /k/, /p/など）は現状の枠組みでは適切な生成が困難であることが予想される。それは、共鳴音と非共鳴

## 研究計画・方法（つづき）

音とでは、その生成メカニズムが構音的に顕著に異なるため、話者性による音の変形の様子が異なるためである。そこで、これらを異なる2ストリームとして扱い、個別に構造構築、空間探索・定位する方法を検討する。更に現状の構造からの音声生成では、与えられた正解音との距離関係のみを用いて次音を定位している。時刻  $t$  の音を定位するために、与えられた正解音と、推定された時刻  $t-1$  までの音全部を用いて、時刻  $t$  の音を定位することも可能である。こうすることで、発声全体の構造的妥当性は向上すると考えられる。音声生成における全体的な構造制約の導入によって、自然性向上を図る。

### 2. 基本周波数などの韻律的特徴を包含した音声の構造的表象の導出 [担当：峯松・修士・博士]

上記の構造分析には音の高さの情報、即ち基本周波数の情報は含まれていない。ここでは基本周波数の情報を如何にして音声の構造的表象の中に組み込むのかについて検討する。既に音楽情報処理において、特徴ベクトル系列となった音楽情報に対して、全てのベクトル間スカラー距離を求める形で一つの（巨大な）構造として音楽を表現する方法が提案されている。音楽である以上、音の高さの情報を組み込むことが必須であるが、ここで採られている音楽の音響的表象は、オクターブ構造を前提とした音楽特有の表象となっている。音声の場合、オクターブ構造は基本的に存在しないため、直接的には上記手法は導入不可能である。ここでは、これら音楽情報処理の先行研究例を参考にしながら、スペクトル情報と基本周波数情報とを構造的に融合することを検討する。この時、基本周波数を一旦推定した後に、その差異（コントラスト）を求めると、基本周波数の推定エラーが必ず混入する。エラーフリーの差異推定方法についても検討する。

### 3. 幼児の音声模倣時の音声資料に対する音響的・構造的分析 [担当：峯松・梶川・修士]

実際の幼児は、親の声の何を真似ているのだろうか？申請者が提唱している音声の構造的表象は、どこまで、発達心理学者の言う「語ゲシュタルト」の物理的解釈として妥当なのか？連携研究者の梶川（玉川大学）は、幼児の言語発達の研究をしており、幼児の音声模倣に対して、実際にデータ収集を行なっている。日本人幼児が無意味日本語単語を聞いた時の模倣の様子、彼らが英語単語（有意味であるが、当然未知語である）を聞いた時の模倣の様子などを実際に収録している。ここでは、梶川の協力を仰ぎ、幼児の音声模倣データを提供して戴く。これらに対して、音響的・構造的分析を行なう。逆に、音声構造論から予想される幼児が模倣し易い／模倣し難い単語を実際に聴取・模倣させ、どのような反応を示すのかについても検討する。これらを通して、音声の構造的表象と語ゲシュタルトとの整合性について考察する。

## 平成 22 年度の研究計画・方法

### 1. 身体（声道）情報を制約条件とした空間探索手法の構築 [担当：峯松・博士・修士]

21 年度までの検討は全て、「構造からの音声生成」という新しい枠組みを、その音響的側面に限って行なう検討である。「語ゲシュタルトに身体性を戻すことで構造が音として定位される」のであれば、声道のサイズ／長さ／形状の情報を音声の構造的表象に戻すことで声を定位するのが、本来追求すべき研究対象のはずである。ここでは党教授（北陸先端大学）が提唱している構音器官の制御モデル、あるいは、前田教授（フランス CNRS/ENST）が提唱している構音器官の制御モデルなどを使い、声道の情報を制約条件として与えることで、構造を声として実現する枠組みについて検討する。当然、声道が非常に短ければ、生成される音は「可愛い」声にならざるを得ない。幼児の音声模倣がこれに相当する。

### 2. 構築したシステムの音声模倣モデルとしての発達心理学的評価 [担当：峯松・梶川・博士・修士]

最終的に構築された「構造からの音声生成系」の「音声模倣系」としての発達心理学的な妥当性を検討する。例えば声道情報を制約条件として構成される「構造からの音声生成系」において、上記した構音器官の制御モデルより得られる「個々の構音器官制御の容易さ／困難さ」までも考慮すると、不完全な構音器官制御に基づく「構造からの音声生成系」が構築される。この「くずれた模倣音声」と「構音器官の制御が完全でない幼児が発声した模倣音声」とが、どの程度類似しているのか、を検討することも可能となるだろう。また、この検討を通して、「構造からの音声生成系」をより人間らしく改良することも可能となる。

- [1] N. Minematsu *et al.*, "A consideration of infants' vocal imitation," LNAI4914, 26–39, Springer, 2008
- [2] D. Saito *et al.*, "Speech generation based on infant-like vocal imitation," INTERSPEECH, 1837–1840, 2008
- [3] Y. Qiao *et al.*, "f-divergence is a generalized invariant measure between distributions," INTERSPEECH, 1349–1352, 2008
- [4] 齋藤大輔他, "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
- [5] 峯松信明他, "多次元音楽としての音声モデリングと音声模倣", AI 学会全国大会講演論文集, 1F2-3, 1–4, 2007
- [6] 峯松信明他, "音声の構造的表象を通して考察する幼児の音声模倣と言語獲得", AI 学会 AI チャレンジ研究会, SIG-Challenge-0624-6, 35–42, 2006
- [7] N. Minematsu *et al.*, "Universal and invariant representation of speech," Proc. Int. Conf. Infant Study, 2006
- [8] S. Asakawa *et al.*, "Multi-stream parameterization for structural speech recognition," ICASSP, 4097–4100, 2008
- [9] 朝川智他, "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", 音響学会論文集, 2-P-3, 113–116, 2008
- [10] 峯松信明, "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

**人権の保護及び法令等の遵守への対応（公募要領9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、次頁以降に記入する研究経費の妥当性・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。

21年度に行なう、1) 共鳴音・非共鳴音を含んだ音声に対する空間探索手法の構築、2) 基本周波数などの韻律的特徴を包含した音声の構造的表象の導出、及び、3) 幼児の音声模倣時の音声資料に対する音響的・構造的な分析、のために音声分析サーバを新規に1台購入し、同時に、作業用端末としてのデスクトップPCを3台購入する。音声の構造的表象に関する検討は従来より進めて来たが、本申請課題は、従来とは異なる観点から構造的表象を応用する試みとなっており（従来は、音声認識目的の構造利用だったため、具象→抽象への変換であったが、今回は、音声生成目的であるので、抽象→具象の変換が必要となる）、新たな計算機資源が必要となっている。上記備品は、それに応えるために購入する必須機器である。

22年度では、4) 身体（声道）情報を制約条件とした空間探索手法の構築、5) 構築したシステムの音声模倣モデルとしての発達心理学的評価、などを研究計画として掲げている。声道情報を制約条件として導入するためには、構音音声合成ソフトウェア（ライブラリ）の利用を考えている。また、幼児の音声模倣時の音声収録、或は、それに類する音声収録を東大でも行なう可能性を考慮し、携帯性の高い音声収録機を計上している。更に、評価実験用にノートPCを複数台購入する予定である。

本申請研究は音声工学に限らず、非常に多くの分野の注目を浴びている。これは、今後も、多分野・異分野の研究者との情報交換を頻繁に行なう必要があることを意味する。21年度、22年度と研究打合せ旅費を比較的多く申請しているのはそのためである。なお、消耗品として音声データベース、信号処理ソフトウェアの購入を予定している。前者は、基本周波数の正解ラベルとしてEGG (Electro Glottograph) データが付与されたデータベースの購入を意図しており、後者は、MATLABなどの信号処理ソフトウェアのライセンス費用である。

以上のように、計上した研究経費はどれも本申請課題に対して不可欠なものであり、経費削減は、本課題の円滑な遂行を妨げる可能性が非常に高いことを追記しておく。

研究機関名 | 東京大学

研究代表者氏名 | 峯松 信明

**挑戦的萌芽－6**  
(金額単位：千円)

設備備品費の明細				消耗品費の明細				
(記入に当たっては、挑戦的萌芽研究 研究計画調書作成・記入要領を参照してください。)				(記入に当たっては、挑戦的萌芽研究 研究計画調書作成・記入要領を参照してください。)				
年度	品名・仕様 (数量×単価) (設置機関)		金額	品名	金額			
2 1	音声分析サーバー・DELL PowerEdge1950(1×750)(東京大学)		750	音声データベース 信号処理ソフトウェア	200			
	作業用デスクトップ PC・DELL XPS 720(3×150)(東京大学)		450		200			
	計		1,200	計	400			
2 2	デジタル音声収録機・マランツ MD671(1×150)(東京大学)		150	構音音声合成ライブラリ 増設メモリ	300			
	評価実験用ノート PC・DELL Precision(3×200)(東京大学)		600		100			
	計		750	計	400			
旅費等の明細 (記入に当たっては、挑戦的萌芽研究 研究計画調書作成・記入要領を参照してください。)								
年度	国内旅費		外国旅費		謝金等		その他	
	事項	金額	事項	金額	事項	金額	事項	金額
2 1	研究打合	250	成果発表	300	研究補助(1人×2月×50千円)	100		
	成果発表	300			専門知識提供(2人×10時間×5千円)	100		
	計	550			計	200		
2 2	研究打合	250	成果発表	300	聴取実験謝金(20人×10時間×1千円)	200	論文投稿料	100
	成果発表	300				論文別刷代	50	
	計	550				計	200	計



**研究費の応募・受入等の状況・エフォート**

本欄は、第2段審査（合議審査）において、「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」及び「新学術領域研究」の領域提案型にあつては、「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）	平成21年度研究経費（期間全体の）	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
<b>【本応募研究課題】</b> 挑戦的萌芽研究 (H21~H22)	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)	15	
特定領域研究（公募班） (H21 ~ H22) 東京大学	構造不変性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	構造表象に基づく音声認識系の構築が目的であり、音声生成系とは異なる技術構築が行なわれている。
若手(S)（一般）(H21 ~ H24) 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)	—	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捉え、処理系を構築する。購入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。

研究機関名 東京大学

研究代表者氏名 峯松 信明

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）	平成 21 年度研究経費（期間全体の）	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)	20	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	音声合成を目的としているが、テキスト処理に関する研究なので、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 (%)	