

平成18年度 学術創成研究費 研究終了報告書 (事後評価用)

平成18年3月31日

ふりがな	たなか ほづみ		②所属研究機関・ 部局・職		東京工業大学・名誉教授			
①研究代表者 氏名	田中穂積							
③研究課題名 (英訳名)	言語理解と行動制御 (Natural language understanding and action control)							
④研究経費 (千円未満切捨)	年度	研究経費 (千円)		使用内訳 (千円)				
		交付額	支出額	設備備品費	消耗品費	旅費	謝金等	その他
	平成13年度	90,000	90,000	52,272	14,247	17,400	3,681	2,402
	平成14年度	90,000	90,000	27,454	25,292	16,959	7,418	12,873
	平成15年度	90,000	90,000	27,667	19,626	24,359	3,293	15,053
	平成16年度	90,000	90,000	53,044		24,431	8,078	4,445
	平成17年度	80,000	80,000	42,350		23,596	10,399	3,654
総計	440,000	440,000						
⑤研究組織 (研究代表者及び研究分担者)								
氏名	所属研究機関・部局・職			現在の専門	役割分担 (研究実施計画に対する分担事項)			
田中 穂積 研究分担者	東京工業大学・名誉教授			自然言語処理	研究の統括と管理, 言語理解の研究			
土屋 俊	千葉大学・文学部・教授			哲学	言語と行動に関する認知理論			
山田 友幸	北海道大学・大学院文学研究科・教授			哲学	言語と行動に関する認知理論			
辻 幸夫	慶応義塾大学・法学部・教授			認知言語学	言語と行動に関する認知理論			
山梨 正明	京都大学・大学院人間・環境学研究所・教授			認知言語学	言語と行動に関する認知理論			
楠見 孝	京都大学・大学院教育学研究科・助教授			認知科学	言語と行動に関する認知理論			
丸山 直子	東京女子大学・文理学部・助教授			日本語学	言語と行動に関する認知理論			
白井 清昭	北陸先端科学技術大学院大学・情報科学研究科・助教授			自然言語処理	言語理解の機構 (言語生成)			
奥村 学	東京工業大学・精密工学研究所・助教授			自然言語処理	言語理解の機構 (意味・文脈理解)			
松本 裕治	奈良先端科学技術大学院大学・情報科学研究科・教授			自然言語処理	言語理解の機構 (構文・意味・文脈理解)			
徳永 健伸	東京工業大学・大学院情報理工学研究科・助教授			自然言語処理	言語理解の機構 (意味・文脈・対話理解)			
石崎 雅人	東京大学・社会情報研究所・助教授			自然言語処理	対話理解の機構 (対話管理)			
飯田 仁	東京工科大学・メディア学部・教授			自然言語処理	対話理解の機構 (対話管理)			
牧野 正三	東北大学・大学院工学研究科・教授			音声言語処理	対話理解の機構 (音声認識, 音声分析)			
河原 達也	京都大学・学術情報メディアセンター・教授			音声言語処理	対話理解の機構 (対話分析, 音声理解)			
古井 貞熙	東京工業大学・大学院情報理工学研究科・教授			音声言語処理	対話理解の機構 (対話分析, 音声理解)			
鹿野 清宏	奈良先端科学技術大学院大学・情報科学研究科・教授			音声言語処理	対話理解の機構 (対話分析, 音声理解)			
北橋 忠宏	関西学院大学・理工学部・教授			計算機科学	ソフトウェアロボットの行動制御			
白井 良明	立命館大学・情報理工学部・教授			画像処理	環境認識とロボットの行動制御理論			
小林 哲則	早稲田大学・理工学部・教授			ロボティクス	環境認識とロボットの行動制御理論			
佐藤 誠	東京工業大学・精密工学研究所・教授			ヒューマンインタフェース	ソフトウェアロボットの構築 (動作機能, 動作の実時間映像化)			
中嶋 正之	東京工業大学・大学院情報理工学研究科・教授			コンピュータグラフィクス	ソフトウェアロボットの構築 (動作機能, 動作の実時間映像化)			
原島 博	東京大学・大学院情報学環・教授			コンピュータグラフィクス	行動機能が豊富な3次元ソフトウェアロボットの構築 (豊富な表情制御)			
広瀬 啓吉	東京大学・大学院新領域創成科学研究科・教授			音声言語処理	行動機能が豊富な3次元ソフトウェアロボットの構築 (音声合成と口唇動作の同期)			
小林 隆夫	東京工業大学・大学院総合理工学研究科・教授			音声言語処理	行動機能が豊富な3次元ソフトウェアロボットの構築 (音声合成と口唇動作の同期)			

⑥当初の研究目的

「言語理解と行動制御」に関しては、古くは言語哲学、最近では認知言語学の分野でいくつかの理論も提案されている。相手の行動を促す発話では、発話の文字通りの意味とは異なる意味や意図を相手に伝えたいことがしばしばある。これは発話行為論とよばれており、トップダウン的な理論である。発話のさまざまな現象を網羅的にカバーした理論であるとは言えない。これに対して本研究では、むしろボトムアップにこの問題にアプローチする。具体的な対話データを分析し、そこから得られた知見、理論を実際に動作する計算モデルに組み込む。それを動作させることによりさらに新しい問題の抽出と解決を行なう、というサイクルを含めて、「言語理解と行動制御」に関する新しい学問分野を構築しようとするものである。上記した研究目標を達成するために、次の4つの研究課題に大別して研究を行なう。

- (1) 言語理解の機構に関する研究
- (2) 行動機能が豊富な3次元ソフトウェアロボットの構築
- (3) 言語と行動に関する理論の構築
- (4) 音声対話によるソフトウェアロボットの行動制御システムの試作

最後の研究課題について説明を補足しておく。上記した他の3つの研究課題の研究成果の評価は必ずしも容易ではない。すべてが相互に関連していることも問題である。我々の行なっている言語理解の機構そのものも十分明らかであるとはいえない。未解明の問題が山積しているのである。そこで本研究では、チューリングテストの考え方を採用し、音声対話によりロボットの行動制御を行なうプロトタイプシステムを実際に構築し動作させることにより、研究成果が妥当であるかどうかを判断し検証しながら研究を進める。研究期間が5年であることを考慮し、プロトタイプシステムで理解の対象とする話し言葉の語彙と文の長さには制限を設ける。

⑦研究成果の概要

研究目的に対する研究成果を必要に応じて図表等を用いながら、具体的に記入してください。

日本語、英語、中国語などの自然言語をコンピュータに理解させる言語理解の研究は、これまで主として書き言葉を対象として、理工学、言語学の立場から研究されてきた。理工学の立場からは、質問応答システム、機械翻訳システム、情報検索システムなどの応用システム構築を巡り多数の研究が行われてきた。しかし、言語理解については未解決な問題が山積しているため、応用システムの構築にあたっては、むしろ困難な言語理解の問題を避けたシステムの構築を目指してきたと言える。一方、言語学の立場でも、主として書き言葉を対象に、言語理解にとって必要な意味の問題は解決困難であるとして、それを意識的に避け、文法論を中核に据えた言語研究が行われてきた。しかし言語理解のためには解決が困難であっても、意味の問題に挑戦しなければならない。そして、話し言葉は書き言葉とは異なる側面も持っているため、コンピュータに話しかけ、コンピュータに仕事をさせることを実現しようとする、書き言葉より話し言葉を対象にした言語理解の研究が必要になる。そこで、本学術創成研究では、話し言葉に焦点を当てた研究を行った。以下ではこれまでに得られた研究成果の概要を順に述べる。

(1) 言語理解の機構に関する研究 (田中, 白井(清), 奥村, 松本, 徳永, 石崎, 飯田)

言語理解システムを構築するためには、理工学、言語学のみならず心理学、哲学の立場からの研究成果を統合することが重要である。そして個々の研究成果の有効性を評価するために、コンピュータ内部(仮想空間)に、人間に近い姿をした3次元のソフトウェアロボットを構築し、それに音声による指令を与え仮想空間内で動作させるシステムのプロトタイプを構築し、研究成果の有効性を検証した。プロトタイプシステムを作成するために、対話により相手に簡単な問題を解かせ、そのときの対話履歴を収集して分析した(2001年度報告書)。分析作業は、心理学、認知科学、計算機科学の分野の研究者が協力して行った。動作指令は音声で与え、ロボットはコンピュータ内部でシミュレートするロボット(ソフトウェアロボット)とした。ソフトウェアロボットにしたことで、複雑な動作指令を言語で与えることが可能になった。そして、ロボットに与える音声による動作指令は、ロボットの行動を動画としてコンピュータの画面上に表示する。指令を与えた人間は、ロボットの画面上の動作を見て、ロボットが言語(指令)を理解したかどうかを判断することができる。

本学術創成研究の成果を最終的にこのプロトタイプシステムに収斂させ、成果を評価するために、コンピュータグラフィックスの研究と言語理解の機構に関する研究を行う研究者とを、密接に連携させた共同研究を行った。ロボットの行動の可視化にとれない、これまでの言語理解の研究ではまったく考慮されていなかった不明確性の問題が重要であることが明らかになった。たとえば「右に進め」というロボットに与える指令に含まれる「右」とはどの程度「右」なのか、原理的には無限に存在する右の位置の可能性の中からもっともらしい位置の一つ決めてロボットは行動する必要がある。この問題は、ポテンシャル関数の概念を用いて解決した(2004年度報告書)。ポテンシャル関数による不明確性解消は、ロボットがある地点からある地点まで障害物を避けながら移動する経路探索の問題にも応用可能である。プロトタイプシステム(K4)に実装し動作させてポテンシャル関数の有効性を確認している。次に「右」とは、指令を出した話者から見て「右」なのか、指令の受け手であるロボットから見た「右」なのか曖昧で、相対的な問題である。この問題は、これまで多くの研究者によって論じられてきたが、それを言語理解の機構に組み込みロボットとして動作させた研究は類がない(2004年度報告書)。どの程度右に進めばよいかについても不明確性の問題が存在する。ロボットの大きさや周囲の状況に依存して相対的に決まる。これは「状況に依存した言語理解」を行うことが必要であることを示唆し、今後さらに研究する必要がある。

ロボットの行動の可視化には、空間的不明確性の他に、動作の多様性に関する不明確性がある。たとえば、「コップを掴め」という指令に含まれるコップの掴み方には無限の方法が存在する。ロボットはそこから妥当な掴み方一つ選んで動作しなければならない。この問題は心理学者の提唱したアフォーダンス理論を用いて解決した(2005年度報告書)。

⑦研究成果の概要 つづき

アフォーダンス理論では、コップ(の形状)が、掴み方を決めると考える。この理論を用いて、コップ、フライパンなどさまざまな形状の物体に対して、ロボットが妥当な掴み方を生成する手法を提案した。掴み方に関する不明確性は、ロボット行動の可視化にともない顕在化し、一見単純そうに見える「右」、「掴む」などの言葉の理解には、これまでの単語辞書に記述されている事柄以上の事実をロボットは知っている必要がある。そこで、行動生成を行う行動記述を含む辞書のあり方について検討し、小規模辞書を作成してプロトタイプシステムで利用した。このような行動記述を含む辞書については、今後さらに研究する必要がある。動作指令の話し言葉について、指示詞の指す物を参照・同定する照応問題は、「状況に依存する言語理解」の典型的な問題である。指示詞に付く助詞の種類に基づく意味を考えない参照物同定法として中心化理論がよく使われてきた。しかし、助詞の種類だけを用了参照物同定法に限界があることは明らかである。中心化理論に違反する会話例を容易に見つけることができる。たとえば、青い玉のある机の前にロボットが立っている状況で、「青い玉をつかめ」「赤い玉のある机の前に行け」「それを床に落とせ」という一連の動作指令に含まれる「それ」は「青い玉」なのか「赤い玉」なのか。中心化理論によれば「赤い玉」と判断されるが、この状況では「青い玉」の方が妥当である。人間は「落とす」という動作の前提として、手に持っているものでなければならない、ということを知っているからである。動作するためには動作の手順があり(これは人工知能の研究ではプランニングと呼ばれている)前提が満たされた動作のみを行うことができる。そして動作後には状況変化が生じ、それがまた次の動作の前提になる。このプランニングの考え方を適用して指示詞の参照するものを同定する新しい方法を開発した。このプランベース照応物同定法によると、従来の方法では不可能であった指示物の同定が可能となる(2001年度報告書)。また人間には簡単であっても、修復文(言い直し文)、繰り返し文の存在は言語処理を著しく困難にする。どこからどこまでが繰り返しか、修復文であるかは状況に依存する。この問題に本格的に対処した研究はこれまでなかったといつてよい。指令文を誤解してロボットが行動したとき、それを検知し、修正する新しい方法を開発した。困難であるとして避けてきたこれらの問題に対処する新しい方法を提案し、それらをプロトタイプシステムに組み込み、動作させたことの意義は大きい。プロトタイプシステムの応用として、料理法の質問に対して料理法を動画として教示するシステムも試作している。言語理解システムの前処理段階では、指令文の単語単位の分かち書きが必要になる。これは形態素解析と呼ばれている。書き言葉に比べて話し言葉の文に対する形態素解析の精度はこれまで極めて不十分であった。そこで、話し言葉文の高精度形態素解析システムを構築し、ツールとして一般に公開することを考えている(2004年度報告書)。この形態素解析システムは音声認識にも応用可能である。

(2) 言語と行動に関する理論の構築(土屋, 山田, 辻, 山梨, 楠見, 丸山)

日本語の話し言葉については、省略が多い、助詞の欠落などの事実が知られている。そこで、言語学的な見地から日本語の話し言葉についてコーパスに基づく詳細な分析、検討を行った。プロトタイプシステムへのフィードバックには到らなかったものの、この研究は話し言葉の日本語文法構築の第一歩となる研究成果であると位置づけることができる。言語と行動については認知言語学の研究がある。認知言語学の立場から言語表現を行動パターンにどうグラウンディングさせるか、身体的な意味論、参照点能力に基づく照応研究を行った。「言語と行動」に関する認知言語学的な研究成果は2001-2005年度報告書に詳しい。この方向の基礎研究は今後も継続して進める必要がある。

心理学の立場からは、「前」、「上方」などの空間表現語の不明確性の問題に対して、その度合いについて被験者による応答実験を行い、人間はこれらの空間表現語をどう理解しているかを検討した(2003-2005年度報告書)。得られた成果は、プロトタイプシステム構築で参考にしていく。人間の空間認知を解明する理論の一部をなす研究成果である。哲学の立場からは、人間の言葉によるコミュニケーション機構の論理的な形式化をおこなっている(2005年度報告書)。

(3) 音声処理の研究(古井, 牧野, 河原, 鹿野, 徳永)

音声認識は対話システムのフロントエンドに位置しているため、一層の精度の向上と頑健性が望まれる。クラスター法とよぶ方法を開発し、それが、さまざまな雑音下でも音声認識の精度を低下させないことを確認した。また多数のクラスターモデルを並列に用いることにより、多様な音響言語モデルの変動にもよく適応することを明らかにした(2002-2005年度報告書)。音声認識の精度をさらに向上させるためには、多様な話し言葉音声のカバーするコーパスが必要になることを明らかにした。本研究で作成したプロトタイプシステムとは別に、音声情報(博物館)案内システムを構築し、頑健な音声認識システムの構築を目指したフィールドテストを2002年から現在まで行った。フィールドテストを通じて、多様な話者からなる対話音声データを大量に収集しており、将来の対規模音声認識システムの構築に役立てようとしている。博物館の案内に関する質問に対する応答文の作成方法についても研究し評価している(2002-2005年度報告書)。情報検索を音声で行なうシステム、信念ネットワークに基づく対話システムを開発し、性能を評価した。お茶くみ等のサービスロボットや身振り、手振りを認識するロボットを開発した。「アー」とか「あー」などのフィラーの存在は話し言葉では避けられない。これらは非文法文とよび、言語処理を著しく困難にすることが知られている。この問題は言語理解の機構に関する研究者と共同して新しい方法を開発しそれをプロトタイプシステムに実装した(2003年度報告書)。

(4) 行動機能が豊富な3次元ソフトウェアロボットの構築および音声対話によるその行動制御システムの試作

(中嶋, 白井(良), 北橋, 小林(哲), 佐藤, 小林(隆), 広瀬, 原島, 徳永)

コンピュータグラフィクス関連の研究分担者の多くはプロトタイプシステムの構築に直接関わっており、空間表現や握持に関する不明確性の問題解決などの研究成果を挙げている。2003年度にはニュートン力学を考慮したプロトタイプを開発している(2003年度報告書)。手足の合計25個の関節モデルをもち、それにより自然で多様な動作をコンピュータ画面に実現することに成功している。この研究グループには音声で対話するハードウェアロボットを対象にした研究も含まれている。前述のプロトタイプシステムと異なり、ここでは3次元物体認識が大きな問題となる。言語理解については、ソフトウェアロボットから得られる知見の多くをそのまま用いることができる。しかし、物体認識は言語理解とは異なる側面を持つ問題である。冷蔵庫の扉を開けて指定の飲み物を取り出すサービスロボットの実現を目標としている。物体認識では物体認識誤りが避けられない。そこで認識誤りを対話により絞り込む方法を提案した。ハードウェアロボットを対象にした別の研究では、あいつち、首振り、視線合わせを行うことが可能なハードウェアロボットを開発し、非言語によるコミュニケーションの果たす役割の重要性を明らかにしている(2004, 2005年度報告書)。非言語によるコミュニケーション機能はソフトウェアロボットにも組み込むべき機能の一つである。対話を行うときに対話相手の表情変化も非言語コミュニケーションの一つとして重要なものの一つであるが、動作するロボットに組み込むためには、時期尚早であり、独自に研究を行った。少数のパラメータで、多種多様な表情を生成する新しい手法を提案した(2003-2005年度報告書)。自然な音声を合成する手法を開発するとともに、概念構造から音声による文章生成を行う方法を提案した。

⑧特記事項

この研究において得られた、独創性・新規性を格段に発展させる結果あるいは可能性、新たな知見、学問的・学術的なインパクト等特記すべき事項があれば記入してください。

音声による指令を与え仮想空間内でロボットを動作させるプロトタイプシステムでは、指令を理解した結果を可視化し映像化して、指令を理解したかどうかを判断することができる。このプロトタイプシステムのロボットは人工知能システムである。このシステムを動作させ評価することは、人工知能が実現したかどうかを判断するための、古典的なチューリングテスト以上に厳しい人工知能の評価になっている。なぜなら、チューリングテストは、壁の向こうに人工知能をおき、それに対して質問し(ゲームの場合には、盤面を与え)、人工知能から返される応答を見て、人工知能が実現しているか否かを判断する。人工知能それ自体は壁の向こうにあり、見えない。一方、本研究で構築したプロトタイプシステムの応答は可視化され、ロボットの動作は、指令を与えた側から見える。動作の可視化は、これまでのチューリングテストでは顕在化しない不明確性の問題などが重要であることを明らかにした。そのことの学問的学術的な意義は大きい。これまで以上に深い意味理解、言語理解に基づく動作が必要になるからである。言語理解には人工知能の基本問題(行動のプランニングの問題、障害物を避けた移動経路発見の問題、物体移動に関するフレーム問題など)が含まれている。いずれも学問的・学術的に意味ある研究課題である。本研究課題に挑戦することにより、学問的・学術的な観点から、人工知能の基本問題に対する新たな問題、新たな知見が得られる可能性がある。

本研究で構築したプロトタイプシステムは、フィラーを含む非文法的な文の処理、誤解の検出と修復などに関する新しい解決法を提案し、それを動作するシステムとして実現した点に独創性と新規性がある。ロボット動作の可視化にあたり、不明確性をポテンシャル関数を用いて解決した研究例はこれまで皆無であるとはいえないが、それをロボットの動作機構に組み込み動作させ評価した意義は大きい。誤解の検出と理解についても独創的な手法を開発している。しかも、単に指令をロボット側が誤解していることを検出するだけでは不十分である。動作したロボットは、誤解に伴う動作のやり直しや修復動作を行わなければならないからである。これは実際にロボットを動作させてみてはじめて見出した新しい知見である。修復動作にともない動作の再プランニングの問題を扱わなければならない。

本研究では人間とロボットの1対1の対話を考えていた。1対1のモデルではなく、学校の授業などに見られる1対多の対話モデルも考えられる。サッカーなどのゲームで、選手同士が対話を行いながら目標に向かって行動する場合などは、それが敵方も漏れる可能性があるので、一方通行の授業モデルなどよりさらに複雑な対話モデルになる。敵方はそれを妨害する行動に出たりすることがあるからである。しかも攻撃、防御には協調動作も必要になる。このような対話モデルに取り組んだ研究は皆無であり今後の重要な研究課題である。

動作辞書の必要性については書き言葉の研究ではこれまでまったく考慮されてこなかった。動作に関する辞書記述の問題は、言語レベル(記号レベル、抽象レベル)の記述を、動作レベル(具体レベル)へと、どうグラウンディングさせるかを定める極めて認知科学的な問題でもある。プロトタイプシステムでは小規模な動作辞書を作成し利用した。そこから得た知見を利用し、動作辞書の記述法についてはさらに検討する必要がある。これは書き言葉の研究からは得られないことであり、重要である。

不明確性の問題を含めて、「状況に依存した言語理解」の問題については今後さらに研究する必要がある。省略語の補強、一文ではなく複数の文章理解ではこの「状況に依存した言語理解」が特に重要になる。本研究ではその一部の問題の解決を試みたに過ぎない。未解決の問題が山積している。最近の言語理解の研究では、困難な問題をできるだけ避けた問題解決法を見出そうとする傾向にあるだけに、その重要性を特に指摘しておきたい。

ロボットとの対話(コミュニケーション)は、ロボットの言語理解能力だけでは不十分であることが知られている。うなずきや、視線あわせなど非言語コミュニケーション機能も重要であることが知られている。このような機能をロボットにもたせた研究成果も特記に値する。実験によれば、こうした機能は、自然な対話を促進することを実証している。

以下では、機械的なハードウェアロボットではなくソフトウェアロボットを対象にした理由を述べる。最近のコンピュータグラフィックス技術の進歩は著しく、ハードウェアロボット以上に柔軟で多様な動作を行うソフトウェアロボットを仮想空間内に構築することを可能にした。機械的な制約のため、ハードウェアロボットと人間との間で交わす対話は、これまでせいぜい単語のレベルに留まっていた。これがソフトウェアロボットを研究対象にした第1の理由である。第2に、複数のロボットを相手にした1対多の対話の研究では、複数のソフトウェアロボットを仮想空間に作り出して複数のロボットが存在する状況を作り出すことが簡単にできるが、ハードウェアロボットではコストの問題が浮上する。またハードウェアロボットが複数になれば、故障と修理の問題が無視できない。最近の複数ロボットを用いた研究は、ソフトウェアロボットを用いるべきである。

本学術創成研究から得た重要な結論の一つとして、本研究課題をさらに発展させた学問・学術をわが国が創成した文理融合の新しい学問・学術として Computational Robotics と命名して推進したい。これまでの説明から明らかのように、それは文理融合の格好の研究課題である。

最後に、本プロジェクトで開発したプロトタイプシステムは、工学的にもさまざまな応用が考えられる。たとえば、エンターテインメントロボット、介護ロボット、サービスロボット、サイン言語による聴覚困難者との対話ロボット、都市空間の地図ナビゲーション、情報家電、CTスキャナーから得た内臓の3次元映像の探索、などである。応用システムが具体化すれば、福祉の分野だけでなく、社会に多大なインパクトをもたらす可能性がある。一例を挙げれば、MRTやCTスキャンの結果を3次元の体内(内臓など)立体構造を取り出す技術がすでに開発されているので、内視鏡を飲まずに体内を調べることが可能になる。映像を写し出すカメラもロボットと考えることができるので、言葉による都市空間のナビゲーションシステムと同様に、3次元の体内を自由に探索し、医師がその映像を見ながら診断を下すことができるからである。これは、現在の医療のあり方を変える可能性があると思われる。

⑨研究成果の発表状況

この研究費による成果の発表に限り、学術誌等に発表した論文（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）の全著者名、論文名、学協会誌名、巻（号）、最初と最後のページ、発表年（西暦）、及び国際会議、学会等における発表状況について、3頁以内に収めて記入してください。

- 阿辺川武, 奥村学, 共起情報及び複数格の組み合わせを考慮した係り受け解析, 自然言語処理, 2006.
- I.R.Lane, T.Kawahara, T.Matsui, and S.Nakamura. Out-of-domain utterance detection using classification confidences of multiple topics. IEEE Trans. Audio, Speech & Language Process., 2006.
- I.R.Lane and T.Kawahara. Verification of speech recognition results incorporating in-domain confidence and discourse coherence measures. IEICE Trans., Vol.E89-D, No.3, 2006.
- C.Troncoso and T.Kawahara. Trigger-based language model adaptation for automatic transcription of panel discussions. IEICE Trans., Vol.E89-D, No.3, 2006.
- Tobias Cincarek, Tomoki Toda, Hiroshi Saruwatari, Kiyohiro Shikano, "Utterance-based Selective Training for the Automatic Creation of Task-Dependent Acoustic Models", IEICE Trans. Information and Systems, Vol.E89-D, No.3, 2006.
- Randy Gomez, Akinobu Lee, Tomoki Toda, Hiroshi Saruwatari, Kiyohiro Shikano, "Improving Rapid Unsupervised Speaker Adaptation based on HMM Sufficient Statistics in Noisy Environments using Multi-template Models", IEICE Trans. Information and Systems, Vol.E89-D, No.3, 2006.
- Shigeki Miyabe, Hiroshi Saruwatari, Kiyohiro Shikano, Yosuke Tatakura, "Interface for Barge-in Free Spoken Dialogue System Using Nullspace Based Sound Field Control and Beamforming", IEICE Trans. Fundamentals, Vol.E89-A, No.3, 2006.
- Jinfu Ni, Hisashi Kawai, Keikichi Hirose, "Constrained tone transformation technique for separation and combination of Mandarin tone and intonation", Journal of Acoustical Society of America, to be published (2006).
- Jinfu Ni, Keikichi Hirose, "Quantitative and structural modeling of voice fundamental frequency contours of speech in Mandarin", Speech Communication, (2006).
- 渡辺美知子, 広瀬啓吉, 伝康晴, 峯松信明, "音声聴取時のフィルターの働き--「エート」による後続句の複雑さ予測-", 日本音響学会誌, (2006).
- 杉森絵里子・楠見 孝 ソースモニタリングエラーにおける質判断と時間判断の検討：時間経過が反応バイアスに及ぼす影響 認知心理学研究, 2, (1), 35-45. 2005.April
- 楠見 孝 認知心理学から見た比喻 日本語学 (明治書院), 24, 5月号, 26-36. 2005.
- 杉森絵里子・楠見 孝, ソースモニタリングエラーにおける質判断と時間判断の検討：時間経過が反応バイアスに及ぼす影響 認知心理学研究, 2, (1), 35-45. 2005.Aug.
- 杉森絵里子・中西政志・米田英嗣・常深浩平・楠見 孝, 反復呈示と二重課題がアウトプットモニタリングに及ぼす影響 心理学研究, 76, 244-251. 2005.Sept.
- 楠見 孝, 物語理解における恐怖の生起メカニズム：怪談とメタファ表現研究, 82, 17-26. 2005.Nov.
- 工藤拓, 松本裕治, 相対的な係りやすさを考慮した日本語係り受け解析モデル, 情報処理学会論文誌, Vol.46, No.4, pp.1082-1092, April 2005.
- 中川哲治, 松本裕治, 単語レベルと文字レベルの情報を用いた中国語・日本語単語分割, 情報処理学会論文誌, Vol.46, No.11, pp.2714-2727, November 2005.
- 秋田祐哉, 河原達也. 話し言葉音声認識のための汎用的な統計的発音変動モデル. 電子情報通信学会論文誌, Vol.J88-DII, No.9, pp.1780-1789, 2005.
- 下岡和也, 内元清貴, 河原達也, 井佐原均. 日本語話し言葉の係り受け解析と文境界推定の相互作用による高精度化. 自然言語処理, Vol.12, No.3, pp.3-17, 2005.
- M.Nishida and T.Kawahara. Speaker model selection based on Bayesian information criterion applied to unsupervised speaker indexing. IEEE Trans. Speech & Audio Process., Vol.13, No.4, pp. 583-592, 2005.
- Sadaoki Furui, "Recent progress in corpus-based spontaneous speech recognition", IEICE Transactions on Information and Systems, vol.E88-D, no.3, pp. 366-375 (2005-3)
- Jen-Tzung Chien and Sadaoki Furui, "Predictive hidden Markov model selection for speech recognition", IEEE Transactions on speech and audio processing, vol.13, no.4, pp.377-387 (2005-7)
- Sadaoki Furui, Masanobu Nakamura, Tomohisa Ichiba and Koji Iwano, "Analysis and recognition of spontaneous speech using Corpus of Spontaneous Japanese", Speech Communication, vol.47, pp.208-219 (2005-9)
- Zhipeng Zhang, Toshiaki Sugimura and Sadaoki Furui, "Tree-structured clustering methods for piecewise linear-transformation-based noise adaptation", IEICE Transactions on Information and Systems, vol.E88-D, no.9, pp.2168-2176 (2005-9)
- 鹿野清宏, "音声情報案内システム「たけまるくん」", 未来材料, 第5巻, 第6号, pp.6-7, 2005.
- 望月 亮, 大久保雅史, 小林哲則, "韻律情報を用いたスペクトル変換方式の検討", 電子情報通信学会論文誌 (D-II), vol.J88-D-II, no.11, pp.2269-2276, Nov. 2005.
- Seungzoo JEONG, Naoki HASHIMOTO, Makoto SATO, "Immersive Multi-Projector Display on Hybrid Screens with Human-Scale Haptic Interface", IEICE TRANSACTIONS on Information and Systems, Vol.E88-D, No. 5, pp.888-893 (2005).
- S.jeong, N.hashimoto, M.sato, "Human-scale multi-modal interaction with virtual human based on force feedback", 2005 亜洲芸術科学学会論文集, vol.1, no.1, pp17-18 (2005).
- N. P. Chandrasiri, T. Naemura and H. Harashima, "System for analyzing real-time facial expression responses of a user", 日本顔学会 2005, Vol.5, No.1, p.147 (2005.9).
- 中洲 俊信, N. P. Chandrasiri, 苗村 健, 原島 博, "対話型似顔絵作成 システム NIGAO ～髪モデルの検討～", 日本顔学会 2005, Vol.5, No.1, pp.185 (2005.9).
- 松岡 薫, N. P. Chandrasiri, 苗村 健, 原島 博, "主成分分析を用いた顔画像処理プラットフォーム FaceComposer における顔空間視覚化機能", 日本顔学会 2005, Vol.5, No.1, pp.177 (2005.9).

- Keikichi Hirose, Kentaro Sato, Yasufumi Asano and Nobuaki Minematsu, "Synthesis of F0 contours using generation process model parameters predicted from unlabeled corpora: Application to emotional speech synthesis", *Speech Communication*, Vol.46, Nos.3-4, pp.385-404 (2005-7). (Special Issue on Quantitative prosody modeling for natural speech description and generation)
- Naotake Niwase, Junichi Yamagishi, Takao Kobayashi, "Human walking motion synthesis with desired pace and stride length based on HSMM", *IEICE Trans. Information and Systems*, vol.E88-D, no.11, pp.2492-2499, 2005, Nov.
- Otani, H. Kusumi, T. Kato, K., Matsuda, K., Kern, R.P. Widner, R. Jr., Ohta, N. Remembering a Nuclear Accident in Japan: Did It Trigger Flashbulb Memories? *Memory*, 13, (1), 6 - 20, 2005, Jan.
- Y. Akita and T. Kawahara. Language model adaptation based on PLSA of topics and speakers for automatic transcription of panel discussions. *IEICE Trans.*, 2005.
- Panikos Heracleous, Satoshi Nakamura, Kiyohiro Shikano, "Simultaneous Recognition of Distant-Talking Speech of Multiple Talkers Based on the 3-D N-best Search", *Journal of VLSI Signal Processing*, Kluwer Academic Publishers, 2005.
- 山梨正明 2004. 「理論言語学の新展開--認知言語学のパラダイム」, 『人工知能学会誌』, Vol.19, No.1, pp.84-94.
- 中本敬子・楠見孝, 比喩材料文の心理的特性と分類: 基準表作成の試み, *読書科学*, 48(1), 1-10. 2004, Sep.
- 浅原正幸, 松本裕治, 日本語固有表現抽出におけるわかち書き問題の解決, *情報処理学会論文誌*, Vol.45, No.5, pp.1442-1450, May 2004.
- 工藤拓, 松本裕治, カーネル法を用いた言語解析における高速化手法, *情報処理学会論文誌*, Vol.45, No.9, pp.2177-2185, September 2004.
- 船越孝太郎, 徳永健伸. 話し言葉における言い直しの処理. *情報処理*. Vol.45. No.10. pp.1032 -- 1037. 2004. Oct
- 石崎雅人, コミュニケーションの科学へ向けて, *人工知能学会*, Vol. 19, No. 4, pp. 468-474, 2004.
- 駒谷和範, 上野晋一, 河原達也, 奥乃博. 音声対話システムにおける適応的な応答生成を行うためのユーザモデル. *電子情報通信学会論文誌*, Vol.J87-D-II, No.10, pp.1921--1928, 2004.
- T.Kawahara, M.Hasegawa, K.Shitaoka, T.Kitade, and H.Nanjo. Automatic indexing of lecture presentations using unsupervised learning of presumed discourse markers. *IEEE Trans. Speech & Audio Process.*, Vol.12, No.4, pp. 409--419, 2004.
- H.Nanjo and T.Kawahara. Language model and speaking rate adaptation for spontaneous presentation speech recognition. *IEEE Trans. Speech & Audio Process.*, Vol.12, No.4, pp. 391--400, 2004.
- 下岡和也, 南條浩輝, 河原達也. 講演の書き起こしに対する統計的手法を用いた文体の整形. *自然言語処理*, Vol.11, No.2, pp.67--83, 2004.
- Chiori Hori and Sadaoki Furui, "Speech summarization: an approach through word extraction and a method for evaluation", *IEICE Transactions on Information and Systems*, vol. E87-D, no. 1, pp. 15-25 (2004).
- Zhipeng Zhang and Sadaoki Furui, "Piecewise-linear transformation-based HMM adaptation for noisy speech", *Speech Communication*, vol. 42, iss.1, pp. 43-58 (2004).
- Koji Iwano, Takahiro Seki and Sadaoki Furui, "Noise robust speech recognition using F0 contour information", *IEICE Transactions on Information and Systems*, vol. E87-D, no. 5, pp. 1102-1109 (2004).
- Sadaoki Furui, Tomonori Kikuchi, Yousuke Shinnaka and Chiori Hori, "Speech-to-text and speech-to-speech summarization of spontaneous speech", *IEEE Transactions on speech and audio processing*, vol. 12, no. 4, pp. 401-408 (2004).
- Takahiro Shinozaki and Sadaoki Furui, "Dynamic Bayesian network-based acoustic model incorporating speaking rate effects", *IEICE Transactions on Information and Systems*, vol. E87-D, no. 10, pp. 2339-2347 (2004).
- Mate Szarvas and Sadaoki Furu, "Phonology and morphology modeling in a very large vocabulary Hungarian dictation system", *IEICE Transactions on Information and Systems*, vol. E87-D, no. 12, pp. 2791-2800 (2004).
- 山出慎吾, 馬場朗, 芳澤伸一, 李晃伸, 猿渡洋, 鹿野清宏, "実環境における頑健な音声認識のための音韻モデルの教師なし話者適応", *電子情報通信学会論文誌*, Vol.J87-D-II, No.4, pp.933--941, April 2004.
- 中島淑貴, 柏岡秀紀, ニックキャンベル, 鹿野清宏, "非可聴つぶやき認識", *電子情報通信学会論文誌*, Vol.J87-D-II, No.9, pp.1757-1764, September, 2004.
- Ryo Mochizuki, Tetsunori Kobayashi, "A low-band spectrum envelope reconstruction method for PSOLA-based F0 modification," *IEICE Transaction of Information and System*, Vol.E87-D, No.10, pp.2426--2429. Oct. 2004.
- Shinya Kiriya, Keikichi Hirose, and Nobuaki Minematsu, "A spoken dialogue system for document information retrieval utilizing topic knowledge." *Systems and Computers in Japan*, John Wiley & Sons, Vol.35, Issue 5, pp.67-82 (2004-5).
- Dhany Arifianto, Tomohiro Tanaka, Takashi Masuko, Takao Kobayashi, "Robust F0 estimation of speech signal using harmonicity measure based on instantaneous frequency," *IEICE Trans. Information and Systems*, vol.E87-D, no.12, pp.2812--2820, 2004, Dec.
- 小林隆夫, 徳田恵一, "コーパスベース音声合成技術の動向[IV] - HMM 音声合成方式 -," *電子情報通信学会誌*, vol.87, no.4, pp.322--327, 2004, Apr.
- 鍛冶伸裕, 黒橋禎夫: 迂言表現と重複表現の認識と言い換え, *自然言語処理*, Vol.11, No.1 (2004.1)
- 山肩洋子, 河原達也, 奥乃博, 美濃導彦. 音声対話システムにおける物体指示のための信念ネットワークを用いた曖昧性の解消. *人工知能学会誌*, Vol.~19, No.~1, pp. 47--56, 2004.
- Rajkishore Prasad, Hiroshi Saruwatari, and Kiyohiro Shikano, "Probability Distribution of Time-Series of Speech Spectral Components," *IEICE Trans. Fundamentals*, March, 2004.
- 西村竜一, 西原洋平, 鶴身玲典, 李晃伸, 猿渡洋, 鹿野清宏, "実環境研究プラットフォームとしての音声情報案内システムの運用", *電子情報通信学会論文誌*, Vol.J87-D-II, No.3, 2004
- 榎原 靖, 白井良明, 島田伸敬 "対話を用いた物体認識のための照明変化への適応", *電子情報通信学会論文誌* Vol.~J87-D-II, No.~2, 2004.
- Shinya Kiriya, Keikichi Hirose, and Nobuaki Minematsu, "A spoken dialogue system for document information retrieval utilizing topic knowledge." *Systems and Computers in Japan*, John Wiley & Sons, 2004.

- 田中穂積, 徳永健伸. ロボットとの会話—人工知能からのアプローチ. 情報処理. Vol.44. No.12. pp.1247 -- 1252. 2003.
- Kuriyama, N., Yamagishi, K., and Kusumi, T. Effects of instructions and representation on mathematical problem solving, *Psychological Reports*, 93, pp.~603--609, 2003.
- Junko Araki, "Action Vectors: Modeling Spatial Relations between Objects and Routes", *Journal of Universal Computer Science*, 9(9), pp. 1046--1072, Sep. 2003
- 中川哲治, 工藤 拓, 松本 裕治, "Support Vector Machine を用いた形態素解析と修正学習法の提案", 情報処理学会論文誌, Vol.44, No.5, pp.1354-1367, 2003.
- 船越孝太郎, 徳永健伸, 田中穂積. 音声対話システムにおける日本語自己修復の処理. 自然言語処理. Vol.10. No.4. pp.33--53. 2003. Jul
- Tetsuro Takahashi, Kozo Nawata, Kentaro Inui, Yuji Matsumoto. Effects of Structural Matching and Paraphrasing in Question Answering. *IEICE Transactions on Information and Systems*, E86-D(9), pp. 1677-1685, Sept., 2003.
- 藤田篤, 乾健太郎. 語彙・構文的言い換えにおける変換誤りの分析. Exploring Transfer Errors in Lexical and Structural Paraphrasing. 情報処理学会論文誌, Vol. 44, No. 11, pp. 2826-2838, 2003.11.
- T.Nishikawa, H.Saruwatari, K.Shikano, "Blind source separation of acoustic signals based on multistage ICA combining frequency-domain ICA and time-domain ICA," *IEICE Trans. Fundamentals*, Vol.E86-A, No.4, pp.846--858, April 2003
- T.Nishikawa, H.Saruwatari, K.Shikano, "Stable learning algorithm for blind separation of temporally correlated acoustic signals combining multistage ICA and Linear Prediction," *IEICE Trans. Fundamentals*, Vol.E86-A, No.8, pp.2028--2036, August, 2003.
- H. Saruwatari, S. Kurita, K. Takeda, F. Itakura, T. Nishikawa, and K. Shikano, "Blind Source Separation Combining Independent Component Analysis and Beamforming," *EURASIP Journal on Applied Signal Processing*, Vol.2003, No.11, pp.1135--1146, 2003.
- 滝澤正夫, 榎原 靖, 白井良明, 島田伸敬, 三浦 純, "サービスロボットのための対話システム", システム制御情報学会誌, Vol.~16, No.~4, pp.~174--182, 2003.
- Piperakis Romanos, Suguru SAITO, Masayuki NAKAJIMA: "Automatic Calibration of Magnetic Motion Trackers using a Bayesian-Neural Cascade", *The Journal of the society for art and Science*, Vol.~2, No.~3, pp.~105--113 (2003)
- Atsuhiko Sakurai, Keikichi Hirose and Nobuaki Minematsu, "Data-driven generation of F0 contours using a superpositional model," *Speech Communication*, Vol.40, No.4, pp.535-549 (2003-6).
- Junichi Yamagishi, Masatsune Tamura, Takashi Masuko, Keiichi Tokuda, Takao Kobayashi. A training method of average voice model for HMM-based speech synthesis. *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, Vo.E86-A, No.8, pp.1956-1963, 2003, Aug.
- Junichi Yamagishi, Masatsune Tamura, Takashi Masuko, Keiichi Tokuda, Takao Kobayashi. A context clustering technique for average voice models. *IEICE Trans. Information and Systems*, Vol.E86-D, No.3, pp.534-542, 2003, Mar.
- 滝澤正夫, 榎原 靖, 白井良明, 島田伸敬, 三浦 純. サービスロボットのための対話システム. システム制御情報学会誌, Vol.~16, No.~4, 2003.
- Atsuhiko Sakurai, Keikichi Hirose and Nobuaki Minematsu, "Data-driven generation of F0 contours using a superpositional model," *Speech Communication* (2003).
- 田中穂積, 徳永健伸. コンピュータが拓く新しい言語世界. 月刊 言語, Vol.~31, No.~3, pp.~16-22, 2002, Mar.
- Yamanashi, Masa-aki. "Cognitive Perspectives on Language Acquisition", in *Studies in Language Science*, No.~2, pp.~107-116, 2002.
- 山梨正明. 「言語科学における記述・説明の妥当性—認知言語学の視点からみた言語研究の展望」, 『日本語文法』, Vol.~2, No.~2, pp.~3-28, 2002.
- 山梨正明. 「焦点連鎖とメンタルパスからみた言語現象」, 玉村文郎(編) 『日本語学と言語学』, 明治書院, pp.~469-484, 2002.
- 山梨正明. 「認知言語学からみた 21 世紀の言語研究」, 『英文学研究』(日本英文学会), Vol.~78, No.~2, pp.243-244, 2002.
- 駒崎久明・楠見 孝. 洞察問題解決における表象変換と制約緩和: マッチ棒代数課題の反復試行による検討 認知科学, 9(2), 274-284, 2002, June.
- Matsuo, M. and Kusumi, T. Salesperson's procedural knowledge, experience and performance: An empirical study in Japan. *European Journal of Marketing*, 36(7&8), 840-854, 2002, Sept.
- 浅原正幸, 松本裕治, 「形態素解析のための拡張統計モデル」, 情報処理学会論文誌, Vol.~43, No.~3, pp.~685-695, March 2002.
- 松本裕治, 「自然言語処理におけるシステム混合法の利用」, 電子情報通信学会論文誌 D-II, 招待論文, Vol.~J85-D-II, No.~5, pp.~709-716, May 2002.
- 浅原正幸, 米田隆一, 山下亜希子, 伝康晴, 松本裕治, 「語長変換を考慮したコーパス管理システム」, 情報処理学会論文誌, Vol.~43, No.~7, pp.~2091-2097, July 2002.
- 鈴木, 牧野: 「離散型 HMnet における最適な状態数の自動決定法」, 電子情報通信学会論文誌, Vol.~J85-D-II, No.~2, pp.193-199. (2002)
- 長野, 鈴木, 牧野: 「HMM を用いた複数 n -gram による言語モデルの構築」, 情報処理学会論文誌, Vol.~43, No.~7, pp.~2075-2081. (2002)
- 伊藤亮介, 駒谷和範, 河原達也. 機器操作マニュアルの知識と構造を利用した音声対話ヘルプシステム. 情報処理学会論文誌, Vol.~43, No.~7, pp.~2147-2154, 2002.
- 長谷川将宏, 秋田祐哉, 河原達也. 談話標識の抽出に基づいた講演音声の自動インデキシング. 情報処理学会論文誌, Vol.~43, No.~7, pp.~2222-2229, 2002.
- 李晃伸, 河原達也, 鹿野清宏. 音素環境独立 HMM を用いた混合ガウス分布選択による音響尤度計算の削減. 情報処理学会論文誌, Vol.~43, No.~7, pp.~2214-2221, 2002.