

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 3 月 31 日現在

機関番号：12601

研究種目：基盤研究（B）

研究期間：2009～2011

課題番号：21300061

研究課題名（和文） 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

研究課題名（英文） Expressive Multi-language Speech Synthesis Based on the Generation Process Model and Its Use for Automatic Speech Translation

研究代表者

広瀬 啓吉 (HIROSE KEIKICHI)

東京大学・大学院情報理工学系研究科・教授

研究者番号：50111472

研究成果の概要（和文）： 基本周波数パターン生成過程モデル (F_0 モデル) の枠組みで、多言語の韻律制御の研究を統合的に進め、 F_0 モデルの指令の差分に着目した韻律の適応手法を開発し、焦点制御、調子の変換、話者変換を実現した。さらに、学習音声コーパス F_0 パターンあるいは生成される F_0 パターンの F_0 モデル近似を行い、HMM 音声合成の品質向上を達成するとともに、焦点付与を達成した。成果に立脚し、談話・意図を再現する音声翻訳の実験を行った。

研究成果の概要（英文）： A unified study on prosody control for multi-languages was conducted based on the generation process model of fundamental frequency contours (F_0 model). We developed a method of prosody adaptation, where differences in F_0 model commands were learned from parallel speech corpus and were applied to baseline speech. Focus control, style conversion and voice conversion were realized. Furthermore, by approximating F_0 contours of training speech corpus and/or generated F_0 contours using the F_0 model, we improved the quality of synthetic speech by the HMM-based speech synthesis. Also, we added focus control. Based on the above results, experiments were conducted on conveying discourse information and intentions in speech Translation.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	4,800,000	1,440,000	6,240,000
2010年度	4,100,000	1,230,000	5,330,000
2011年度	4,600,000	1,380,000	5,980,000
年度			
年度			
総計	13,500,000	4,050,000	17,550,000

研究分野：情報科学

科研費の分科・細目：情報学 ・ 知覚情報処理・知能ロボティクス

キーワード：生成過程モデル, 基本周波数パターン, コーパスベース韻律制御, 音声自動翻訳, 談話焦点, HMM 音声合成, 声質と調子, 音声モーフィング

1. 研究開始当初の背景

コーパスベース波形接続手法の導入によって、合成音声の品質は格段に向上したが、それは、多量の音声コーパスを前提としたものであり、合成される音声も、主として平坦な読み上げ朗読調に限定されている。Nick

Campbell (当時 ATR) らは、自然な場面での音声コーパスを収集することによって、コーパスベース波形接続手法の枠組みで“表現力豊かな音声合成”を目指したが、音声コーパス収集の困難さから思ったような成果が得られなかった。コーパスベース波形接続手

法の問題点は、合成する音声の声質・調子の音声コーパスを多量に用意することであり、これは、朗読音声であれば可能ではあるが、例えば、感情音声などの様々な調子を安定に保つことは困難である。“表現力豊かな音声合成”を達成するためには、声質や調子の柔軟な制御が可能な音声合成手法を開発することが必須となっている。

近年、音声認識で用いられている隠れマルコフモデル (HMM) を音声合成に用いる HMM 音声合成が開発されている。これは、従来のソース・フィルタモデルに基づく分析合成系であるが、音響パラメータを統計的に表現するため、分析合成系で問題となっていた、パラメータの選択と接続という問題を、尤度最大の基準で自動的に解決するという利点がある。このため、必要となる音声コーパスを大幅に削減することが可能となり、また、HMM の適応などによりごく少量の音声コーパスから種々の声質・調子の音声へ変換することも試みられている。このため、声質や調子の柔軟な制御が可能な音声合成手法として、研究者の注目を得ている。

HMM 音声合成の大きな問題点として、その韻律の表現が挙げられる。一般に音声認識では、韻律的特徴は用いられていないので問題となっていないが、音声合成では、韻律をモデルに組み込む必要がある。特に基本周波数 (F_0) の表現が問題となっている。現在は、単純に各フレームでの F_0 値 (と有声/無声情報) を対象としているが、 F_0 に代表される韻律的特徴は、超分節の特徴ともいわれるように、単語、句といった長時間に渡る特徴で、フレーム単位での表現にはそぐわない。また、 F_0 の時間パターン (F_0 パターン) は、本来、発話の言語情報や構造と密接に関係しているが、HMM 音声合成で得られる F_0 パターンからそれを読み取ることは困難である。このため、我々の韻律に関する従来の知見を導入することが困難で、極度に不自然な韻律を生成する可能性があった。また、同一の発話内容でも調子が異なると発話構造は同一とは限らず、このような場合、HMM の適応は良好に機能しない。

このような観点から、本研究では、我々が従来から進めている F_0 パターン生成過程モデル (F_0 モデル) の枠組みを導入することにより、韻律の観点から HMM 合成音声の品質向上と意図・態度・感情などの柔軟な韻律制御を達成するものである。

韻律には、日本語の単語アクセント、中国語の声調、英語の強勢などのように言語固有性が強い部分と、統語に対応したイントネーション、疑問調での文末上昇イントネーション、韻律句内での発話速度変化 (句末音節の伸長など) などのように言語間の差異が小さい部分がある。このため、一言語のみを対象

として研究を進めるよりも、言語間の異同に着目して研究することにより、見通しの良い韻律制御手法を達成することが期待される。従来、韻律の研究は、個々の言語に限定して行うことが一般的であったが、このような観点から、ここでは、日本語、中国語を中心として、多言語の韻律を統合的に研究し、音声合成のための韻律制御手法を開発する。 F_0 モデルは、対数 F_0 パターンを、句頭から句末に向かう緩やかな起伏のフレーズ成分に、局所的な起伏のアクセント成分 (中国語では声調成分) が重畳したもとして表現し、それぞれの成分が、それぞれ、離散的なインパルス、ステップ関数で記述される指令によって生成されるとしたもので (図 1)、フレーズ成分は、統語情報等と対応し、言語の差が小さいのに対し、アクセント成分 (声調成分) は、単語のアクセント型、音節の声調型等と関連し、言語差が大きいといった、言語間の異同を明確に表現することが可能で、研究に最適な韻律モデルとなっている。

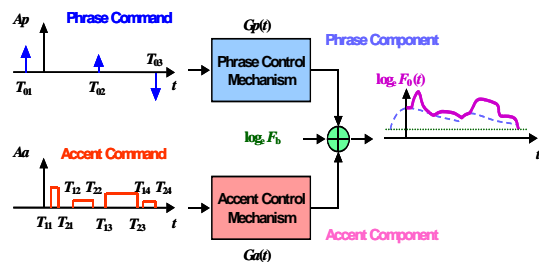


図 1. F_0 モデル

社会の国際化に伴い、言語の壁を意識しないコミュニケーションを実現する音声翻訳が注目されているが、それは、表層の言語情報の伝達のみである。本研究は、意図、態度、感情、個性 (声質) といったパラ言語・非言語情報を合成音声で表現する枠組みを、多言語に対して実現することを目指すものとなっており、その成果を音声自動翻訳に適用し、話者の意図・態度・感情、個性の伝達も可能とする枠組みを構築することを視野に入れている。

2. 研究の目的

現在のテキスト音声合成は、高品質の合成音声を得られるものの、ほぼ平坦な朗読調に限られている。音声で伝達される情報は、語義といった基本的な言語情報に止まらず、談話の焦点といった、高次の言語情報、あるいは、意図、態度、感情、個性といったパラ・非言語情報があり、円滑なマンマシンコミュニケーションの実現の観点からは、これらの情報を適切に表現し得る音声合成手法の開発が急務となっている。このためには、少量の音声コーパスにより (理想的には音声コー

パスによらずに) 種々の声質・調子の音声合成することが可能な“柔軟な音声合成手法”の確立が前提となる。この観点から HMM 音声合成が、期待されているが、その韻律表現に問題があることが指摘されている。この観点から、本研究では、我々が開発し、その有効性が実証されている F_0 モデルを前提とした韻律制御手法を開発することにより、HMM 音声合成ベースとした柔軟な高品質音声合成を達成することを目的とする。 F_0 モデルは、言語・パラ/非言語情報に対応した F_0 パターンの概形制御が可能という特長を有し、韻律的特徴の柔軟な制御が達成できる。

本研究の、顕著な特徴は、日本語、中国語等の多言語に渡り、統合的に韻律を研究し、各言語の韻律制御を達成することにある。

社会の国際化に伴い、音声コミュニケーションを言語の違いを意識することなく行い得る音声自動翻訳に対するニーズが高まっている。近年、音声自動翻訳の性能は格段に向上したが、そこでは、言語内容の伝達のみが対象とされ、発話者の意図・態度・感情、あるいは個人性の伝達は考慮されていないという問題がある。ここでは、デフォルトの話者の朗読調音声合成システムを各言語について構築した上で、本研究で開発する手法により、発話者の声質・調子あるいは意図・態度・感情等を翻訳後の音声に反映させることを行う。この過程において、もとの発話の意図・態度・感情等を自動的に検出することも必要となるが、それについては、研究対象としない。

3. 研究の方法

既に、 F_0 モデルの指令を、入力テキストの言語情報から推定し、 F_0 パターンを生成して、HMM 音声合成のそれと置き換えて合成する手法を日本語で開発している。この手法では、指令に変更を加えることで、声質・調子の柔軟な制御が可能な音声合成が可能である。

本研究では、まず、中国語等の音声の韻律的特徴を、 F_0 モデルを用いて分析し、談話情報といった高次の言語情報、意図・態度・感情といったパラ/非言語情報との関係を明らかにした上で、日本語のそれとの異同を定量的に記述する。さらに、得られた結果に基づき、各言語に最適な韻律生成の枠組みを構築した上で、テキストの言語情報を入力として、 F_0 モデルの枠組みで、所与の調子の韻律を生成する手法を開発する。コーパスベース手法を基本とするが、アクセント/声調成分と比べ、高次の言語情報と深く関係し、広い時間範囲の特徴であるフレーズ成分については、ルールベースの導入を試みる。これによって言語情報等との対応関係が明確で、言語適応性が高い韻律制御手法を開発する。

次に、任意の話者について指定した調子の

音声を、少量の音声コーパスで生成する手法を開発する。これは、基準とするナレータの音素モデル (HMM) の適応と調子を考慮した韻律的特徴の変換を行うものである。韻律の変換については、変換前と変換後の F_0 モデルの指令の差分を学習し、それを変換前の音声に適用することを行う。さらに、話者の発話について全体的な声質と調子を捉えるとともに、談話情報や意図・態度・感情といった発話毎に動的に変化する情報を抽出する手法を開発する。これらの成果を統合し、もとの発話の声質・調子、あるいは意図・態度・感情等を翻訳後の音声に反映させることを行う。

4. 研究成果

(1) 韻律の言語間比較と中国語の韻律制御

日本語、中国語に加え、英語、ドイツ語、ポーランド語、ベンガル語について、その韻律的特徴を分析、比較した。その結果、ベンガル語では、各句頭に位置するアクセント指令が、ほぼ常時、負の値を取ることを見出した。一般的に言って、アクセント/声調成分は言語による差が大きいものの、フレーズ成分は、統語構造、句の長さとの対応の観点からみると、言語によらず、下記のように共通点が多いことを明らかにした。

- 深い統語構造で大きなフレーズ指令が生起する傾向にある。
- 隣り合うフレーズ指令の間隔には上限がある。

また、中国語については、他言語と比べ、フレーズ指令の頻度が高いことを確認した。これは、中国語では、第三声など、 F_0 がフレーズ成分より大幅に下降することが多く、他言語と比べ、フレーズ成分が、ある程度の値を常時保つ必要があるためと考えられる。この結果から、日本語の知見を基に、中国語について、ルールベースで F_0 モデルのフレーズ指令を与えるフレーズ成分制御手法を開発した。さらに、声調成分については、各音節の F_0 パターンのうち、前後の音節の声調の影響を受けず、安定した形状を示す声調核の F_0 パターンを蓄積して選択し、 F_0 レンジ、音節長に合わせて伸縮して直線補間するコーパスベース手法を開発した。フレーズ成分制御手法と統合し、HMM 音声合成と組み合わせる音声合成手法を開発した。合成音声の聴取により、従来の HMM 音声合成と比較した優位性を確認した。開発した韻律制御手法では、まず、フレーズ成分を生成し、次にフレーズ指令の情報をを用いて声調核 F_0 パターンの選択と伸縮を行う。中国語では、強調が置かれた語句等に必ずフレーズ指令が置かれることを利用し、簡単な制御により焦点付与した音声の合成を行い、その有用性を示した。また、声調核 F_0 パターンに着目した韻律変換手法

を開発し、感情音声の合成を行った。

(2) F_0 モデルの指令の差分に着目した韻律の適応

声質や調子の変換前後の同一言語内容の音声で、 F_0 モデル指令の1対1対応が得られる時、指令の差分に着目する韻律変換手法を提案し、まず、合成音声に焦点を付与することを行った。焦点付与では、 F_0 モデル指令の1対1対応が得られる場合が多いからである。HMM 音声合成の学習音声コーパスとして用いている ATR503 文から 50 文を選択して、音声コーパスと異なる女性話者 1 名が発声した、焦点を付与した場合と付与しない場合の音声（パラレル音声）について、 F_0 モデルの指令の差分を、言語情報と焦点の位置情報を入力とする 2 分木用いて学習し、焦点を付与しない音声から付与した音声を合成した。差分を内挿、外挿する音声モーフィングを行い、聴取実験により手法の有効性を確認した。この方法では、パラレル音声コーパスの発話者は、ベースラインの（焦点なしの）音声合成用音声コーパスのそれと同じである必要がないという利点がある。

同様の手法により、調子の変換を行うことも進めた。調子としては、丁寧とぞんざいの 2 つを選択して音声合成を行い、聴取実験により評価した。その結果、丁寧では手法が有効であるが、ぞんざいでは効果が限られることが明らかとなった。ぞんざいでは変換前後の指令の対応についての考察が必要である。

声質（話者）変換は、変換前後の声質のパラレル音声コーパスから、音声特徴量の Gaussian Mixture Model (GMM) を構築し、変換前の声質の音声を入力として、声質を変換するものである。発話の言語情報を使わずに変換を行うことが前提となる。 F_0 について、単に平均値と分散のみの変換しか行っておらず、各アクセント成分、フレーズ成分の強弱といった制御が行われていないという問題点がある。ここではパラレル音声コーパスについて、 F_0 モデルの指令を抽出し、前期と同様の手法で、指令の差分を学習し、それによって、韻律の変換を行う手法を開発した。従来手法で変換した音声と提案手法で変換した音声の比較聴取（どちらが目標声質に近い）を行った。“+1:提案手法、-1:従来手法、0:差がない”の評価基準で、 0.419 ± 0.09 となり、提案手法の優位性が示された。

(3) F_0 モデル指令自動抽出手法の改良

観測あるいは合成で生成される F_0 パターンについて、 F_0 モデルの近似を行い、指令パラメータを自動抽出する新しい手法を開発した。従来手法では、 F_0 パターンを補間して連続的なパターンを得た後、その微分の極値から、まずアクセント指令の位置と大きさ

を求め、次に、得られるアクセント成分を F_0 パターンから引いて得られる残差パターンからフレーズ指令を推定していたが、局所的な F_0 の起伏等により思ったような精度が得られなかった。また、抽出されたアクセント成分の誤差により、残差パターンがフレーズ成分に対応したものにならないという問題点もあった。ここでは、母音部の F_0 パターンのみからフレーズ指令/成分を推定し、その後、アクセント指令/成分を推定する手法を開発した。この方法では、各モーラについて、言語情報から得られる高 F_0 レベル・低 F_0 レベル情報を用い、フレーズ指令/アクセント指令の位置と大きさを推定する。HMM 音声合成との親和性を考慮し、HMM の分布統合で利用する言語情報を用いた。人間の視察による指令抽出を念頭に置いた手法で、表 1 の様に、従来手法と比較し、大幅な性能向上を得た。ただし、recall, precision は下記で定義される指標である。

$$\text{recall (R)} = \frac{\text{\#of commands correctly extracted}}{\text{total \#of commands}}$$

$$\text{precision (P)} = \frac{\text{\#of commands correctly extracted}}{\text{total \#of extracted commands}}$$

表 1. F_0 モデルの指令抽出精度の比較 (%)

	Conventional		Proposed	
	R	P	R	P
Phrase com.	93.3	88.8	96.7	80.4
Accent com. (onset)	83.4	94.0	91.6	97.9
Accent com. (reset)	78.5	88.5	89.2	95.4

従来手法でも、言語情報を制約とすることで、精度が向上するが、提案手法の性能は得られない。

(4) HMM 音声合成の韻律制御の改良

HMM 音声合成で生成される F_0 パターンについて、 F_0 モデルを用いた 2 つの改良を行った。1 つは、音声コーパスの F_0 パターンに対し、 F_0 モデルによる近似と補間を行うもので、もう 1 つは、生成された F_0 パターンに対し、 F_0 モデルの最良近似を行うものである。

F_0 は有声区間では、連続的な値を持つが、無声区間では値を持たない。そのままでは、これを HMM で表現することは出来ず、HMM 音声合成では、有声区間の連続値 HMM と無声区間の離散値 HMM を切り替える Multi-space probability distribution HMM (MSD-HMM) が一般的に用いられている。しかしながら、学習コーパスの有声/無声の誤りや F_0 抽出誤り、有声/無声の切り替えと Δ 特徴とのズレ、等により、音声合成で不自然な F_0 パターンを生成する結果となっていた。特に、中

国語は、第三声など、 F_0 抽出誤り、有声/無声判定誤りが多く発生するため、大きな問題となっている。ここでは、HMM 音声合成の学習コーパスの F_0 パターンについて、(3)により、 F_0 モデルの指令を抽出し、 F_0 モデルによる近似・補間したものとするを行った。当然、連続 F_0 の HMM の枠組みで合成を行う。従来手法 (MSD-HMM) よりも原音声に近いパターンが得られることを示した。なお、有声/無声の切り替えは、音素に従って行った。日本語では、 F_0 抽出誤り、有声/無声判定誤りが、中国語より少なく、それによる音質の劣化は小さいと考えられるが、不自然な F_0 の動きの生成を抑えることが音質の向上に役立つと考えられる。合成音声の比較聴取を行い、(2)と同様の手法 (+1~-1) で評価したところ 0.143 ± 0.124 となり手法の有効性が示された。なお、日本語では、有声/無声の切り替えは、MSD-HMM によっており、中国語の様な音素情報による切り替えを行うと音質は若干劣化する。今後、音声コーパスの F_0 パターンの F_0 モデル近似精度との関係を、調べる必要がある。当然、提案手法で生成される F_0 パターンは F_0 モデルの枠組みに近いものとなり、“柔軟な音声合成”の韻律制御に適していると考えられる。

HMM 音声合成で生成される F_0 パターンは、不自然な F_0 の動きを有することがあり、また言語情報、パラ・非言語情報との明示的な対応が得られず、変換を行うことが容易でない。この F_0 パターンを、 F_0 モデルで近似したものに修正することにより、音質の向上と、韻律制御の柔軟性を付与することが出来る。 F_0 モデルの近似は、(3)の F_0 モデル指令自動抽出手法を用いて行う。修正前後の F_0 パターンで音声合成を行い、比較聴取したところ、(2)と同様の評価で 0.126 ± 0.083 を得た。

F_0 モデルによる近似を行うことで、事後に例えば(2)のように指令を修正することが可能となり、柔軟な韻律制御が達成される。ここでは、焦点付与を行い、聴取実験によりその有効性を確認した。

(5) 音声翻訳の実験

上記の音声合成の適応の成果もとに、談話・意図を再現する音声翻訳の実験を行った。

5. 主な発表論文等

[雑誌論文] (計 4 件)

1. Qinghua Sun, Keikichi Hirose, and Nobuaki Minematsu, "A method for generation of Mandarin F_0 contours based on tone nucleus model and superpositional model," *Speech Communication*, Vol.54, Issue 8, pp.932-945 (2012-10). 査読有
<http://www.sciencedirect.com/science/journal/01676393>

2. Tatsuya Matsuda, Keikichi Hirose, and Nobuaki Minematsu, "HMM-based F_0 contour synthesis using the generation process model," *Acoustical Science and Technology*, Acoustical Society of Japan, 採録決定 (2012-7). 査読有
https://www.jstage.jst.go.jp/browse/ast/33/0/_contents
<http://journals.acoustics.jp/ast-archive/>
3. Miaomiao Wang, Keikichi Hirose, and Nobuaki Minematsu, "Improvement of prosody in HMM-based speech synthesis using generation process model," *Journal of Signal Processing*, vol.15, no.4, pp.279-282 (2011-7). 査読有
<http://www.risp.jp/Product.html>
4. Tetsuya Matsuda, Keikichi Hirose and Nobuaki Minematsu, "HMM-based synthesis of fundamental frequency contours using the generation process model," *Journal of Signal Processing*, vol.14, no.4 pp.277-280 (2010-7). 査読有
<http://www.risp.jp/Product.html>

[学会発表] (計 13 件) (論文集有の国際会議の主なもので 4, 11 以外は査読有)

1. Hiroya Hashimoto, Keikichi Hirose, and Nobuaki Minematsu, "Improved automatic extraction of generation process model commands and its use for generating fundamental frequency contours for training HMM-based speech synthesis," *Proceedings INTERSPEECH*, Portland, September 9-13, accepted (2012-9).
2. Miaomiao Wang, Miaomiao Wen, Keikichi Hirose, and Nobuaki Minematsu, "Emotional voice conversion for mandarin using tone nucleus model – small corpus and high efficiency," *Proceedings of International Conference on Speech Prosody*, Shanghai, May 22-25, pp.163-166 (2012-5). Student paper award (2nd prize)
3. Keikichi Hirose, Hiroya Hashimoto, Jun Ikeshima, and Nobuaki Minematsu, "Fundamental frequency contour reshaping in HMM-based speech synthesis and realization of prosodic focus using generation process model," *Proceedings of International Conference on Speech Prosody*, Shanghai, May 22-25, pp.171-174 (2012-5).
4. Keikichi Hirose, "Fundamental frequency contour generation process model for improved and flexible control of prosodic features in hmm-based speech synthesis," *Proceedings of International Symposium on Frontiers of Research on Speech and Music*,

- Gurgaon, January 18-20, pp.302-306 (2012-1). (Invited Speech)
5. Keikichi Hirose, Tatsuya Matsuda, Hiroya Hashimoto, and Nobuaki Minematsu, "Representing fundamental frequency contours generated by hmm-based speech synthesis using generation process model," *Proceedings of IEEE International Workshop on Machine Learning for Signal Processing*, Beijing, September 18-21, 6 pages in CD-ROM proceedings (2011-9).
 6. Keikichi Hirose, Keiko Ochi, Ryusuke Mihara, Hiroya Hashimoto, Daisuke Saito, and Nobuaki Minematsu, "Adaptation of prosody in speech synthesis by changing command values of the generation process model of fundamental frequency," *Proceedings INTERSPEECH*, Florence, August 28-31, pp.2793-2796 (2011-8).
 7. Miaomiao Wen, Miaomiao Wang, Keikichi Hirose, and Nobuaki Minematsu, "Prosody conversion for emotional Mandarin speech synthesis using the tone nucleus model," *Proceedings INTERSPEECH*, Florence, August 28-31, pp.2797-2800 (2011-8).
 8. Shyamal Das Mandal, Anal Haque Warsi, Tulika Basu, Keikichi Hirose, and Hiroya Fujisaki, "Analysis and Synthesis of F_0 Contours for Bangla Readout Speech," *Proceedings Oriental COCOSDA*, Kathmandu, November 24-25, 6 pages in CD-ROM proceedings (2010-11).
 9. Keikichi Hirose, Keiko Ochi, and Nobuaki Minematsu, "Control of prosodic features in corpus-based generation of fundamental frequency contours based on the generation process model," *Proceedings IEEE International Conference on Signal Processing*, Beijing, pp.629-632 (2010-10).
 10. Miaomiao Wang, Miaomiao Wen, Keikichi Hirose, and Nobuaki Minematsu, "Improved generation of fundamental frequency in HMM-based speech synthesis using generation process model," *Proceedings INTERSPEECH*, Makuhari, September 26-30, pp.2166-2169 (2010-9).
 11. Keikichi Hirose, Keiko Ochi, Miaomiao Wang, Tatsuya Matsuda, Miaomiao Wen, and Nobuaki Minematsu, "Using F_0 contour generation process model for improved and flexible control of prosodic features in HMM-based speech synthesis," *Proceedings of 21st Conference on Electronic Speech Signal Processing*, Berlin, September 8-10, pp.84-93 (2010-9). (Invited Speech)
 12. Miaomiao Wang, Keikichi Hirose, and Nobuaki Minematsu, "Generation of

fundamental frequency in HMM-based TTS using generation process model," *Proceedings of International Conference on Speech Prosody*, Chicago, May 11-14, 4 pages in CD-ROM Proceedings (2010-5).

13. Keiko Ochi, Keikichi Hirose, and Nobuaki Minematsu, "Control of prosodic focus in corpus-based generation of fundamental frequency contours of Japanese based on the generation process model," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, April 20-April 24, pp.4485-4488 (2009-4).

〔図書〕 (計 2 件)

1. Keikichi Hirose, "Prosodic corpora based on F_0 contour generation model and automatic extraction of model parameters," *Computer Processing of Asian Spoken languages*, Edited by S. Itahashi & C. Tseng, Consideration Books, Los Angeles, pp.180-183 (2010-3).
2. Keikichi Hirose, Qinghua Sun "On the prosodic features for emotional speech," *Frontiers in Phonetics and Speech Science*, Edited by G. Fant et al., The Commercial Press, Beijing, pp.263-274 (2009-12).

〔産業財産権〕

- 出願状況 (計 0 件)
- 取得状況 (計 0 件)

〔その他〕

ホームページ等 (研究業績)

<http://www.gavo.t.u-tokyo.ac.jp/~hirose/cv/curriculumvitae.pdf>

6. 研究組織

(1) 研究代表者

広瀬 啓吉 (東京大学・情報理工学系研究科・教授)

研究者番号 : 5 0 1 1 1 4 7 2

(2) 研究分担者

峯松 信明 (東京大学・情報理工学系研究科・准教授)

研究者番号 : 9 0 2 7 3 3 3 3

(3) 連携研究者

河合 剛 (北海道大学・メディア・コミュニケーション研究院・准教授)

研究者番号 : 7 0 3 1 2 9 8 1