

TR-IT-0211

音声基本周波数制御の研究
A study on fundamental frequency
control

平井 俊男
Toshio Hirai

1997, 3

概要

このテクニカル・レポートでは、1993年4月から1997年3月までの4年間、筆者がATR音声翻訳通信研究所で行った音声の基本周波数制御に関する研究を扱う。(以下、基本周波数を F_0 と省略する。) これらの研究を分類すると以下の通りとなる:

- 統計的手法を用いた F_0 制御規則の自動抽出.
- 韻律境界の自動検出.
- 複数の音声データベース間の F_0 制御規則の比較.

Abstract

This report describes the studies of fundamental frequency control in Japanese (fundamental frequency is abbreviated to F_0). They are classified into three categories:

- Automatic extraction of F_0 control rules using statistical analysis.
- Automatic detection of prosodic boundaries.
- Comparison of F_0 control rules derived from multiple speech databases.

目次

1	4年間の研究のあらまし	1
2	各研究カテゴリの概要	3
2.1	統計的手法を用いた F_0 制御規則の自動抽出	3
2.2	韻律境界の自動検出	3
2.3	複数の音声データ・ベース間の F_0 制御規則の比較	3
3	謝辞	4
4	付録: 発表文献	6

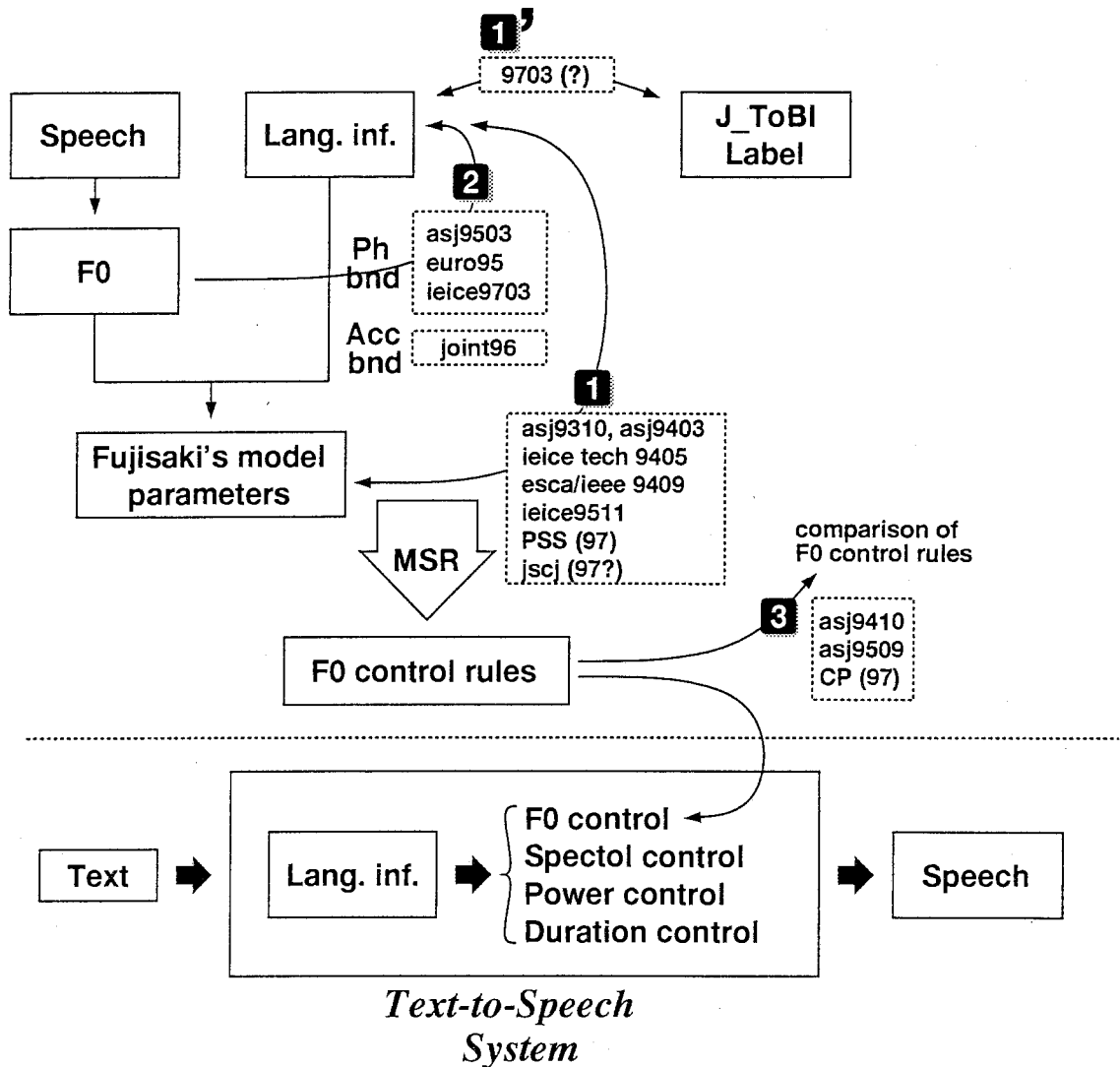


Fig. 1: Research area block diagram.

1 4年間の研究のあらまし

1993年4月から1997年3月までの4年間に、筆者がATR出向中に扱った音声基本周波数(以下、 F_0 と略記)の制御に関する研究を、発表文献の略号を使ってFig.1に示す。また、略記された文献名を以下に示す:

asj: 音響学会講演論文集

CP: "Computing Prosody" from Springer

euro: Eurospeech

esca/ieee: ESCA (European Speech Communication Association)/IEEE workshop on Speech Synthesis

ieice: 電子情報通信学会論文誌

ieice tech: 電子情報通信学会音声研究会

joint: ASA & ASJ joint meeting

jscj: Journal of Systems and Computers in Japan

PSS: "Progress in Speech Synthesis" from Springer

これらの文献を 4. 付録: 発表文献 に添付した. これらの研究は, 以下の 3 つのカテゴリに分類できる:

- ❶ 統計的手法を用いた F_0 制御規則の自動抽出.
- ❷ 韻律境界の自動検出.
- ❸ 複数の音声データベース間の F_0 制御規則の比較.

次章で各研究カテゴリの概要について説明する.

2 各研究カテゴリーの概要

2.1 統計的手法を用いた F_0 制御規則の自動抽出

音声データの F_0 パターンを、藤崎らによって提案されたモデル [1, 2, 3] (以下、藤崎モデル) で分析し、得られる藤崎モデル・パラメータの値とその発話に含まれる言語情報 (アクセント句のアクセント型やフレーズのモーラ数等) との関係を図らによって提案された空間多重分割型数量化法 (Multiple Split Regression) [4, 5] による統計的分析を施して、言語情報から藤崎モデル・パラメータの値を推定する規則を自動抽出する。得られた規則は F_0 に関するものであるため、 F_0 制御規則と捉えられる。抽出された F_0 制御規則は、規則合成装置の F_0 制御部として用いることが可能である。すなわち、入力文の言語情報を F_0 制御規則に与えることにより藤崎モデル・パラメータが推定され、これをもとに合成音声の F_0 パターンが計算される。

asj9503 から jscj までの研究では、筆者が独自に構築した韻律ラベルにより上述の分析を行ったが、汎用的な韻律ラベル・システム J.ToBI のラベルを用いて F_0 制御規則を自動抽出する研究も手掛け、投稿を検討している。投稿予定の原稿も 4. 付録: 発表文献 に収めた。

2.2 韻律境界の自動検出

前節で扱った手法では、言語情報と藤崎モデル・パラメータの値との関係を調べる必要があり、言語情報の付与は人手を要することから、言語情報を音声データから自動的に得ることが可能となれば、 F_0 制御規則生成の完全自動化に一步近づくこととなる。言語情報のうち、最も基本的な韻律情報として、韻律境界の情報がある。これらは発話音声の韻律的な切れ目の情報であり、区切られた区間は言語情報の付与単位となる。このカテゴリーの研究では、韻律境界の自動検出を検討した。韻律境界には大きく分けて、韻律句境界とアクセント句境界の 2 種類がある。asj9503, euro95, ic9703 は韻律句境界を、joint96 はアクセント句境界を扱っている。

2.3 複数の音声データ・ベース間の F_0 制御規則の比較

2.1 節で説明した研究では一話者の読み上げ文を分析の対象としたが、さらに別の話者や発話速度を変えて収録した音声データ・ベースをもとに抽出した F_0 制御規則の傾向を調べたのが、このカテゴリーに分類される研究である。各データ・ベースでの F_0 制御の差を調べることで、以下のことが期待できる:

- F_0 制御規則に現れるパラメータ値のうち、個人や発話様式などによって大きく変わるもの、変わらないものが判明する。変化の大きなパラメータは個人性や発話様式の特徴をより多く含んでいると考えられるので、そのパラメータ値を重点的に調べる必要最小量の音声収録用テキストを設計でき、音声収録時に話者の負担を軽減できる。
- 複数の F_0 制御規則の内分点、外分点を求め、それを規則合成装置で用いることにより、新たな F_0 制御規則が得られる (新しい規則の安定性については別途検討が必要)。

現時点では、扱ったデータ・ベースの数が少ないので十分な検討をしたとまでは言えないが、幾つかの規則のパラメータ値の傾向は、話者や発話速度によっては大きく変化しないことが確認できた。

3 謝辞

ATR 音声翻訳通信研究所での研究の機会を与えて下さった山崎 泰弘社長, 匂坂 芳典室長, 樋口 宜男室長, 出向元オービス総研の関連部署の皆様へ深謝します。研究時, 数多くの有益なコメントを頂いた匂坂室長, 樋口室長, Nick Campbell 主幹研究員に感謝します。

岩橋 直人元第二研究室研究員によって提案された統計分析手法は, F_0 制御規則の自動抽出に関する研究で重要な位置を占めている。また, 深田 俊明研究員, 渡辺 秀行研究員 (以上, 第一研究室), 加藤 宏明研究員 (ATR 人間情報通信研究所) とのディスカッションによって, それぞれ「韻律句境界の自動検出での実験方法」「統計的分析手法一般」「知覚実験での被験者の評価の分析」において, 有意義な示唆を与えて頂いた。さらに, 橋本 誠研究員, 伊藤 嘉治研究員, 藤沢 謙研究員, 北川 敏研究員 (以上, 第二研究室), 第一研究室の皆様にも, 研究に関して有意義なコメントを頂いた。ボストン大学での共同研究 (1996, Jan.-Aug.) では, Mari Ostendorf 教授に研究面での重要な示唆を与えて頂いた。また, 同教授のゼミの皆様には, 研究のみならず, ボストンでの生活のアドバイスを数多く頂いた。皆様へ感謝します。

研究に用いるプログラムやデータ・ベースを作成・維持・管理して下さったエンジニア, ラベラーの皆様へ感謝します。研究時に発生する事務的な作業をサポートして下さった企画課, 秘書の皆様へ感謝します。

研究には直接関与しなかったが, 両親, 姉夫婦をはじめとする, 精神的な面で研究を支えて下さった人々に感謝します。

参考文献

- [1] H. Fujisaki and H. Sudo. Synthesis by rule of prosodic features of connected Japanese. In *7th International Congress on Acoustics*, pp. 133-136, 1971.
- [2] H. Fujisaki and K. Hirose. Analysis of voice fundamental frequency contours for declarative sentence of Japanese. *J. Acoust. Soc. Jpn (E)*, Vol. 5, No. 4, pp. 233-242, Oct. 1984.
- [3] 広瀬啓吉, 藤崎博也, 河井恒, 山口幹雄. 基本周波数パターン生成過程モデルに基づく文章音声の合成. *信学論*, Vol. J72-A, No. 1, pp. 32-40, Jan. 1989.
- [4] 岩橋直人, 匂坂芳典. 空間分割型数量化法による音声制御の統計モデリング. *日本音響学会講演論文集*, pp. 237-238, Oct. 1992.
- [5] N. Iwahashi and Y. Sagisaka. Duration modelling with multiple split regression. In *Proc. Eurospeech*, pp. 329-332, 1993.

4 付録: 発表文献

学会誌論文

- A-2 平井 俊男, 岩橋 直人, 樋口 宜男, 匂坂 芳典: “統計的手法を用いた基本周波数制御規則の自動抽出,” 電子情報通信学会論文誌, Vol.J78-D-II, No.11, pp.1572-1580 (1995, 11)
- A-11 平井, 樋口, 匂坂: “韻律制御パラメータ数を考慮した評価尺度に基づく韻律句境界の自動検出,” 電子情報通信学会論文誌, Vol.J80-D-II, No.2, pp.513-520 (1997, 2)

書籍

- B-1 T. Hirai, N. Iwahashi, N. Higuchi, and Y. Sagisaka: “Automatic extraction of F_0 control rules using statistical analysis” in “Progress in Speech Synthesis” edited by J.P.H. Van Santen, J. Olive, J. Hirschberg, and R. Sproat, pp.333-346, Springer (1997, 2)
- B-15 T. Hirai, N. Higuchi, and Y. Sagisaka: “Comparison of F_0 control rules derived from multiple speech databases” in “Computing Prosody” edited by N. Campbell, N. Higuchi, and Y. Sagisaka, pp.211-223, Springer (1997, 2)

国際会議等

- C-1 T. Hirai, N. Iwahashi, N. Higuchi, and Y. Sagisaka: “Automatic Extraction of F_0 Control Parameters Using Statistical Analysis,” Proceedings of 2nd ESCA (European Speech Communication Association)/IEEE workshop on Speech Synthesis, pp.57-60 (1994, 9)
- C-5 T. Hirai, N. Higuchi, and Y. Sagisaka: “Automatic detection of major phrase boundaries using statistical properties of superpositional F_0 control model parameters,” Proceedings of EUROSPEECH95, pp.1341-1344 (1995, 9)
- C-9 T. Hirai, M. Ostendorf, N. Higuchi, and Y. Sagisaka: “Automatic detection of accentual phrase boundaries and modeling of prosodic features,” Proceedings of 3rd Joint Meeting of Acoustical Society of America and Acoustical Society of Japan, pp.1205-1210 (1996, 12) (Abstract: Journal of Acoustic Society of America, Vol.100, No.4, Pt.2, p.2848, 5pSC2)

研究会

- D-1 平井, 岩橋, 樋口, 匂坂: “統計的手法を用いた基本周波数パターン制御指令の自動分類,” 電子情報通信学会技術報告, SP94-12, pp.47-54 (1994, 5)

学会全国大会

- E-1 平井, 岩橋, H. Valbret, 樋口, 匂坂: “統計的手法による基本周波数パターンの制御,” 日本音響学会研究発表会講演論文集, pp.225-226 (1993, 10)
- E-3 平井, 岩橋, 樋口, 匂坂: “統計的手法を用いた基本周波数制御モデルパラメータの自動抽出,” 日本音響学会研究発表会講演論文集, pp.223-224 (1994, 3)
- E-5 平井, 岩橋, 樋口, 匂坂: “種々の音声コーパスから自動生成された F_0 制御規則の違いについて,” 日本音響学会研究発表会講演論文集, pp.271-272 (1994, 10-11)
- E-7 平井, 樋口, 匂坂: “臨界制動モデルパラメータの分布特性を用いたフレーズ境界の自動推定のための尺度の検討,” 日本音響学会春季研究発表会講演論文集, pp.315-316 (1995, 3)
- E-9 平井, 匂坂, 樋口: “複数話者の音声データベースから自動生成された F_0 制御規則の分析,” 日本音響学会秋季研究発表会講演論文集, pp.209-210 (1995, 9)

投稿予定

- X-1 平井, 樋口: “韻律ラベリング・システム J.ToBI を用いた基本周波数制御規則の自動抽出”