

基本周波数パターン生成過程モデルに基づく 日本語学習者音声の韻律の分析

平野 宏子[†] 顧文涛^{††} 広瀬 啓吉^{††} 峯松 信明[†] 河合 剛^{†††}

[†] 東京大学大学院 新領域創成科学研究科 〒277-8562 千葉県柏市柏の葉 5-1-5

^{††} 東京大学大学院 情報理工学系研究科 〒113-003 東京都文京区本郷 7-3-1

^{†††} 北海道大学 言語文化部 〒060-0817 北海道札幌市北区北 17 条西 8 丁目

E-mail: [†]{hiran,wtgu,hirose,mine}@gavo.t.u-tokyo.ac.jp, ^{††}goh@kawai.com

あらまし 日本語学習者が自然な発音・イントネーションで話したいと願うとき、ピッチ・アクセント言語の日本語において、その自然性に主要に関わる基本周波数の適切な制御が重要になる。本研究では、母語話者と学習者（標準中国語話者）の基本周波数パターンを比較し、その特徴を分析する。その際、1) 表層の音声に現れる基本周波数パターンの特徴を、 F_0 の局所的な上昇下降の形状や発話全体のレンジの変化の様相から述べ、2) 音声成分分野に広く用いられている基本周波数生成過程モデルを、第二言語の韻律習得の分析に応用し、いくつかのパラメータから個々の特徴を分離して定量的に調べる。モデルによる分析からは以下のことが示された。1) 基底周波数は学習者の方が高い、2) フレーズ指令の生起数が多く、形成される韻律句が短い、3) 文節中にアクセント指令が多く生起し、文節がいくつかの韻律語に分解される、4) 文（節）末で急激な F_0 下降が生じ、局所的に負のアクセント指令が導入される。これらの特徴は、音節毎に声調型を持つ中国語音声の影響、第二言語発話の不慣れ、適切な韻律指導の不足に起因すると考えられる。キーワード 日本語音声教育、標準中国語話者、文朗読音声、 F_0 パターン、基本周波数パターン生成過程モデル

Analysis of prosodic features in native and non-native Japanese using generation process model of fundamental frequency contours

H. HIRANO[†], W. GU^{††}, K. HIROSE^{††}, N. MINEMATSU[†], and G. KAWAI^{†††}

[†] Graduate School of Frontier Sciences, University of Tokyo

5-1-5, Kashiwanoha, Kashiwa-shi, Chiba, 277-8562 Japan

^{††} Graduate School of Information Science and Technology, University of Tokyo

7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033 Japan

^{†††} Institute of Language and Culture Studies, Hokkaido University

Kita 17, Nishi 8, Kita-ku, Sapporo, 060-0817 Japan

E-mail: [†]{hiran,wtgu,hirose,mine}@gavo.t.u-tokyo.ac.jp, ^{††}goh@kawai.com

Abstract We compared F_0 patterns of the utterances by native and non-native (Mandarin Chinese) speakers of Japanese. Model-based analysis indicates that non-native speakers tend to transfer the prosodic characteristics of their native language into the utterances of Japanese. They show the following peculiarities as compared with native speakers: 1) a higher baseline frequency; 2) more frequent occurrences of phrase commands and hence shorter prosodic phrases; 3) a tendency of decomposition of a bunsetsu into multiple prosodic words, and hence redundant addition or incorrect location of accent commands; 4) occasional rapider F_0 falling (mostly at the end of utterance) and hence introduction of negative accent commands there. These observed peculiarities can be ascribed to interference by tonal and syllable-timed natures of Mandarin Chinese, difficulty in prosodic planning during speaking second language, and lack of adequate teaching of the skills for sentence intonation.

Key words L2 learning of Japanese, sentence reading, F_0 patterns, generation process model of F_0 contours

1. はじめに

1.1 音声の韻律的特徴

我々が発する音声は、分節的特徴と超分節的特徴とに分けることができる。分節的特徴とは、音素として個々に区別された音として規定される特徴で、発話内容のテキストと密接に対応するものであり、音韻的特徴とも呼ばれる。一方、超分節的特徴とは、アクセント、イントネーション、リズム、ポーズのような単音にまたがって現れる音声の特徴で、韻律的特徴とも呼ばれる。

韻律的特徴は、知覚面から、声の高さ(ピッチ)、強さ、長さといった特徴を有し、対応する物理量として、声帯振動の基本周波数(F_0)、音源の強度、音韻の継続時間長が挙げられる。

F_0 パターンは、日本語を含む多くの言語の抑揚を表し、文の統語構造や意味の伝達に重要な役割を果たすが、語のアクセントを高低の変化によって表現するピッチ・アクセント言語の日本語では、特にこの F_0 パターンが発話の自然性、日本語らしさに主要に関与する。

1.2 第二言語としての日本語

日本語を第二言語として学ぶ学習者は年々増加傾向にあり、在日留学生数は現在 12 万人を突破している [1]。近年、学習者層や学習環境も多様化し、特に日本に暮らす学習者にとって、日本語らしい自然な発音・イントネーションで話すことは重要な学習目標である [2]。

しかし、学習者が学ぶ日本語教育機関では、現状、音声指導を計画的に行っているところが少ない [3]。その指導は、発話の明瞭さに関わる単音が主で、自然性に関わる韻律にはほとんど時間が割かれていない。

言語情報・パラ言語情報・非言語情報など多くを伝える韻律は、その不適切な使用によって、コミュニケーションを阻害したり、相手の感情を害したり、自己の能力を低く見られるといった、学習者の意図しない不利益を生ずる場合があり、教師も韻律の重要性を認識しつつあるが、時間的余裕がない、指導法が分からない、教材がない、などといった理由から、体系的な指導に至っていない [4]。

このような状況は、日本語教育に関わる研究分野において、単音の習得(清濁の区別や特殊拍の習得)や語アクセントの知覚といった側面への研究アプローチは多いが、句や文レベルにわたる韻律に関する研究が十分に行われていない、といったことも反映しているように思われる。

1.3 研究課題

本研究は、 F_0 パターンを中心に、学習者音声の有する韻律的特徴を母語話者と比較して音響分析し、客観的・定量的に記述する手法を提案する。Hirano ら [5] [6] では、1) 生成された F_0 パターンから観察する手法について述べたが、本報告では、2) F_0 パターンの生成過程モデルを用いた分析 [7] について述べる。現在、国内の日本語学習者の 66 % を占める中国語話者の発話について分析しており、本研究ではその結果について述べる。

以降、第 2 章で作成した発話コーパス、第 3 章で表層の F_0 パターンの概観、第 4 章で F_0 パターン生成過程モデルを用い

表 1 母語話者と中国語話者各 10 名の 7 文朗読による平均モーラ数、平均モーラ長 [sec]、平均発話長 [sec] と正規化発話長。

	中国語話者	母語話者
平均モーラ数	25	
標準偏差	(4.30)	
平均モーラ長 [sec]	0.17	0.12
標準偏差	(0.06)	(0.04)
平均発話長 [sec]	5.01	3.59
標準偏差	(1.26)	(1.17)
最大正規化発話長	1.52	1.28
最小正規化発話長	0.72	0.77
標準偏差	(0.16)	(0.10)

た分析、第 5 章でまとめと今後の課題について述べる。

2. 発話コーパスの作成

2.1 音声収録

中国語話者 30 名、日本語話者 10 名に原稿朗読を依頼し、4m² のスタジオで卓上マイクロフォン (Sony D-38B 型) を使用し、音声を DAT (Sony TCD-D10 PRO) に収録した (48kHz, 16bit)。読み間違いや言い淀みは監督者の指示により文頭から読み直した。また、被験者自身も発話が適切と感じるまで自由に読み直しを行った。原稿は、ルビ付き漢字かな混じりの日本語文で、1) 挨拶表現 21 文、2) 11 ~ 40 モーラの短文 27 文、3) 555 モーラの長文 1 文の 3 セクションを、学習者の既習語彙であること、発音特徴、語句のアクセント核位置を考慮して作成した。文中に読点は含まず、話者が自由に発話した。

収録に際し、日本語学習歴、日本滞在歴、出身地域、民族、学校や家庭で使用していた方言などについてアンケート調査も行った。

2.2 分析に用いた音声

コーパスに含まれる中国語話者の出身地は様々で、家庭で各々の方言を使用している。中国語は方言が多様で、語彙面でも音声面でも異なる特徴を有することが知られている。本研究では、学校でも家庭でも標準中国語を話す、東北・華北出身者 (漢族) 10 名と日本語の東京方言話者 10 名の、25 モーラ程度の短文 7 文 (セクション 2) 発話を中心に分析結果を述べる。中国語話者の日本語学習レベルは中級前半から上級前半、在日期間は 8 ヶ月から 1 年 6 ヶ月である。

2.3 F_0 抽出と音素ラベリング

音声分析ソフト PRAAT [8] を用いて 10ms 毎に F_0 抽出を行った。また、波形、スペクトログラム、 F_0 を視察、音声を聴取しながら、手作業で音素ラベリングを行い、その後 F_0 パターンの様子を観察した。

2.4 文発話長

表 1 に中国語話者と母語話者、各 10 名の 7 文朗読による平均モーラ数、平均モーラ長 [sec]、平均発話長 [sec] とそれらの標準偏差、正規化発話長に関する記述統計を示す。

平均モーラ長は、発話区間中の休止を除去して算出した。発話長は、録音時間から文頭と文末の無音区間を除いた時間長を指す。正規化発話長とは、文毎に算出した 10 名の発話長の中

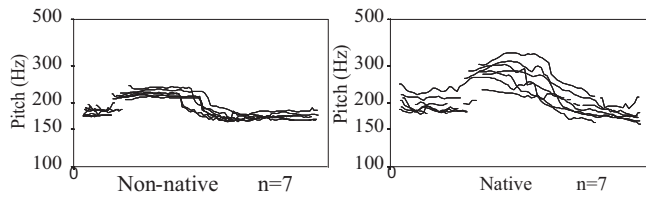


図 1 中国話者 (左) と母語話者 (右) の文章中出现した 7 文節「北海道 { は, の, を, に }」の F_0 パターン。

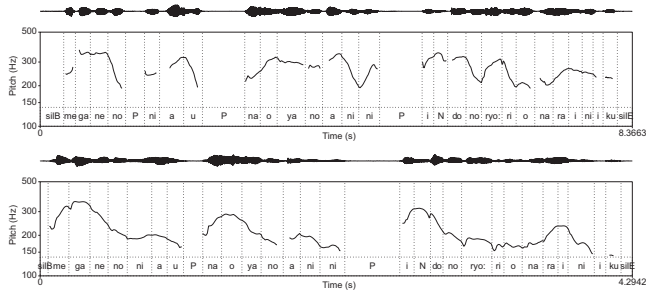


図 2 中国話者 (上) と母語話者 (下) 各 1 名による 1 文朗読音声の F_0 パターン。「眼鏡の/似合う/直也の/兄に/インドの/料理を/習いに/行く」の 8 文節から成る。

央値を、話者毎の発話長で割った値を指す。70 発話で正規化発話長の最大、最小、標準偏差を示した。

母語話者と中国話者では、中国話者の方が全体的に話速が遅く、個人間の差異が大きい。発話長は分散分析の結果、集団間の平均値に統計的有意差が見られた ($F(1,138)=103.37, p<.01$)。

3. 母語話者と中国話者の F_0 パターン

3.1 F_0 パターンの形状とレンジ変化

図 1 は、中国話者と母語話者各 1 名が「北海道観光」に関する 5 分程度の文章を朗読している際に現れた「北海道」+助詞「{ は, の, を, に }」からなる文節 7 つの F_0 パターンを重ねてプロットしたものである。縦軸は基本周波数 (Hz)、横軸は個別の F_0 パターンを重ねて表示するために正規化した時間長である。これは予備実験からのデータであるが、本実験でも見られた中国話者と母語話者の F_0 パターンの特徴がよく現れているので本稿に示す。

母語話者は、文章中文脈や統語情報の違いから、個々の文節を、様々なレンジの F_0 パターンで表現しているが、中国話者は、文章中連続していない 7 箇所の文節から抽出した F_0 パターンであるにも関わらず、高さやレンジが一定であるのが分かる。さらに、ここでの発話「ホッカイドウ (+ 助詞)」は、LHHLLL(L) のアクセント型を有するが、実際の F_0 パターンの形状は、母語話者ではなだらかな丘状の曲線を描く。それに対し、中国話者では、ステップ状の形状が現れている。

図 2 には、本研究の分析文に現れた中国話者と母語話者の 1 文全体の F_0 の変化パターンを示す。母語話者はなだらかな「へ」字型の曲線が大小現れ、文全体のレンジが広いのに対し、中国話者は、母語話者とは異なる、直線的で急峻な F_0 の変動が局所に見られ、レンジが狭いのが分かる。

F_0 レンジ	中国話者	日本語話者
句頭	狭い	広い
後続句	広い	狭い
句末・文末	広い	狭い
F_0 パターン	中国話者	日本語話者
形状	ステップ状	丘状
バリエーション	画一的	多様
移行	急峻	なだらかな
レンジ	狭い	広い

表 2 中国話者と母語話者の F_0 パターンの特徴

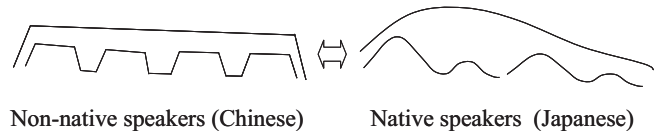


図 3 中国話者 (左) と母語話者 (右) の F_0 パターンの特徴の図式化。

母語話者と中国話者の F_0 パターンの特徴を表 2 にまとめる。また、 F_0 パターンの特徴を大まかに図式化したものを図 3 に示す。

3.2 F_0 パターン分析の課題点と基本周波数パターン生成過程モデルの応用

これまでの F_0 パターンの観測から、学習者の発話には母語話者とは異なる特徴が観察された。ただ、学習者の音声には、語句のアクセント核位置の誤り、局所的な F_0 の上昇下降の様相、全体の F_0 の変化パターンの様相など、母語話者と区別される特徴が層になって出現し、日本語韻律の習得状況によって、話者毎あるいは発話毎に現れる特徴も様々である。そのため、表層の F_0 パターンから、包括的に、定量的にその特徴を記述するのは難しい側面もある。これらの特徴を分離して定量的に記述できる手法が必要である。

そこで、音声合成の分野において、日本語を初めとする多くの言語の F_0 パターンの記述に提案され、広く利用されている基本周波数パターン生成過程モデル [9] を、第二言語の韻律習得に適用して分析に用いることとする。

4. 基本周波数パターン生成過程モデル

4.1 F_0 モデルとは

基本周波数パターン生成過程モデル (以下、 F_0 モデル) は、Fujisaki ら [9] によって提案された数学的モデルで、 F_0 パターンを生成する過程の声帯振動制御機構を、咽頭の生理的・物理的特性に基づいて定量的にモデル化したものである。アクセント指令・フレーズ指令といった少数のパラメータから実測の F_0 パターンに極めてよく近似するパターンを生成することが知られている。

F_0 モデルは、対数軸上で表現した F_0 パターンを、ベースライン成分 (基底周波数)、フレーズ成分、アクセント成分との和として表現したものである (図 4)。

フレーズ成分は、インパルス状のフレーズ指令に対するフレーズ制御機構の応答として表現され、これは句頭から句末にかけて緩やかに下降する成分に対応する。アクセント成分は、

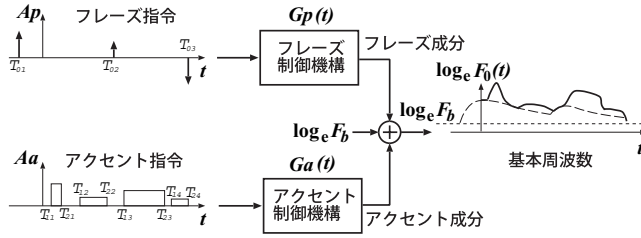


図 4 日本語の基本周波数パターン生成過程モデル

ステップ状のアクセント指令に対するアクセント制御機構の応答として表現され、これは個々の単語または単語の連鎖に付属する局所的な起伏に対応する。両制御機構は、臨界制動 2 次線形系で近似する。これを数式で表現したものが以下である。

$$\log_e F_0(t) = \log_e F_b + \sum_{i=1}^I A_{p_i} G_p(t - T_{0_i}) + \sum_{j=1}^J A_{a_j} \{G_a(t - T_{1_j}) - G_a(t - T_{2_j})\} \quad (1)$$

ここで、式内の変数は以下の意味を持つ。

- F_b : F0 パターンの基底周波数の基底値
- I : 発話中のフレーズ指令の数
- J : 発話中のアクセント指令の数
- A_{p_i} : 第 i 番目のフレーズ指令 (インパルス) の大きさ
- A_{a_j} : 第 j 番目のアクセント指令 (ステップ) の振幅
- T_{0_i} : 第 i 番目のフレーズ指令の生起時点
- T_{1_j} : 第 j 番目のアクセント指令の立ち上がり時点
- T_{2_j} : 第 j 番目のアクセント指令の立下り時点

また、 $G_p(t)$ はフレーズ成分で、次式 (2)

$$G_p(t) = \begin{cases} \alpha^2 t \exp(-\alpha t) & (t \geq 0) \\ 0 & (t < 0) \end{cases} \quad (2)$$

で表され、 $G_a(t)$ はアクセント成分で、次式 (3)

$$G_a(t) = \begin{cases} \min[1 - (1 + \beta t) \exp(-\beta t), \gamma] & (t \geq 0) \\ 0 & (t < 0) \end{cases} \quad (3)$$

で表される。ここで、式 (2)、(3) の α 、 β はそれぞれフレーズ指令、アクセント指令に対する系の速さを定める固有角周波数であり、 γ はアクセント成分の上限を定める相対飽和値である。これらは話者毎、発話毎、言語差による変動が比較的小さいため、通常 $\alpha = 3/s$ 、 $\beta = 20/s$ 、 $\gamma = 0.9$ が用いられる。また、フレーズ指令には、式 (3) に影響しない要素として、負のフレーズ指令を持ち、文の終わり等に生起して、その時点まで継続するフレーズ成分をリセットする作用を持つ。

4.2 パラメータの抽出

成澤ら [10] によって F_0 モデルパラメータの自動抽出法が提

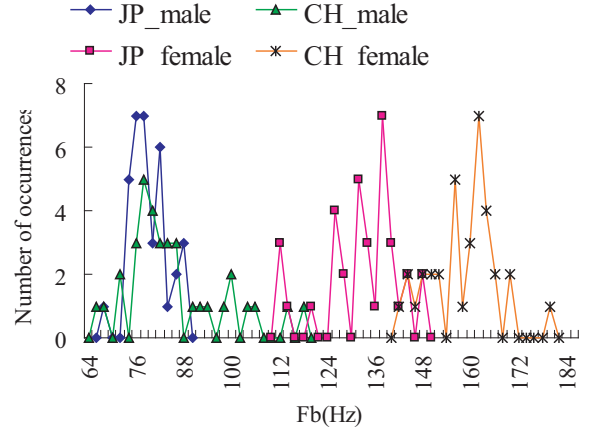


図 5 中国語話者、母語話者、男女別の基底周波数 F_b (Hz) の分布。各 5 話者 $\times$ 7 文発話。

表 3 中国語話者、母語話者、男女別の基底周波数 F_b (Hz) の平均値と標準偏差。各 5 話者 $\times$ 7 文発話。

	CH_male	CH_female	JP_male	JP_female
mean(Hz)	84.5	157.3	78.1	131.7
SD	12.1	8.6	4.6	9.8

案されているが、学習者音声には母語話者には見られない種々の特徴が存在するため、パラメータの自動抽出はかなり困難である。そこで、自動抽出によってあらかじめ得られた全 140 発話 (中国語話者、母語話者、各 10 名 $\times$ 7 発話) の個々のパラメータから、手作業によって修正を施したものを最終的なパラメータ値として分析に用いた。

その際、 F_0 は、音声ファイルを 10kHz、16bit でダウンサンプリングし、遅れ時間比例窓長の自己相関関数を用いた手法 [11] によって 10ms 毎に抽出した。発話全体で得られた F_0 パターンと自動抽出によるパラメータに対し、PRAAT [8] を用いて行った手作業の音素ラベリング情報、音声波形、言語情報と対応を取りながら、基底周波数、フレーズ指令、アクセント指令の大きさと位置を最終的に決定した。なお、フレーズ指令、アクセント指令の固有角周波数 α 、 β 、アクセント成分の相対飽和値 γ はデフォルトの 3/s、20/s、0.9 を統一して用いた。本報告では、基底周波数、フレーズ指令・アクセント指令の生起数に関して分析した結果を述べる。

4.3 基底周波数 F_b

母語話者と学習者の各文発話における基底周波数 F_b を図 5 に示す。表 3 に平均値と標準偏差を示す。各グループ、5 話者 $\times$ 7 文 = 35 発話からデータを得た。

F_b は、母語話者より中国語話者の方が高いことが示された。特に女性の F_b は、母語話者 (131.7Hz) に比べ、中国語話者 (157.3Hz) がかなり高い。男性では、 F_b の分布のピークはどちらのグループも 78Hz 辺りに位置しているが、中国語話者では、 F_b の高い発話は 100Hz を超え、母語話者の女性の値と重なるほどに達しており、平均は中国語話者の方が高くなっている。

表 4 文発話に生起する中国語話者，母語話者，男女毎のフレーズ指令の数の平均値と標準偏差．各 5 話者 × 7 文発話．

	CH_male	CH_female	JP_male	JP_female
mean	4.69	4.83	2.69	3.40
SD	1.45	1.22	0.80	0.65

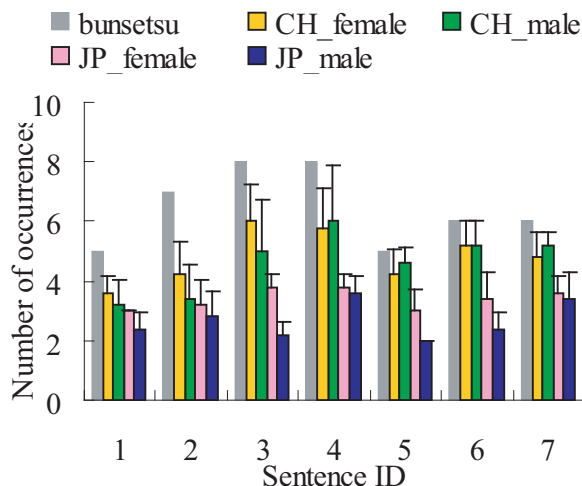


図 6 7 文発話に出現するフレーズ指令数の平均値．4 グループ (中国語話者女性・男性，母語話者女性・男性) 各 5 名による発話．棒グラフ上のバーは標準偏差．グラフ左側の 'bunsetsu' は各文中の文節数．

4.4 フレーズ指令

中国語話者，母語話者の各文発話中に生起するフレーズ指令の数の平均値と標準偏差を表 4 に，男女別に示す．図 6 に各 7 文について，4 グループの指令の生起数を示す．フレーズ指令は発話中，文節境界のいずれかの場所に，それにやや先行して生起する．各文中，左側の縦棒は，文が有する文節数を参照として載せたものである．

母語話者の文中に生起するフレーズ指令の数は，男性で平均 2.69，女性 3.40 に対し，中国語話者は男性 4.69，女性 4.83 と頻度が高い．図 6 中，5 文，6 文の標準偏差のバーを見ると，中には文節数と同じだけフレーズ指令が立つものもあることが分かる．文節毎にフレーズを立て直している例である．

4.5 アクセント指令

中国語話者，母語話者の各文中発話に生起するアクセント指令の数の平均と標準偏差を表 5 に，男女別に示す．図 7 に文毎の内訳を示す．図 7 の各文中，左側の縦棒は文全体のモーラ数，次の左の縦棒は文節数を示す．アクセント指令の生起数は，母語話者では文節数と同じかやや多いくらいだが，中国語話者は，より頻度が高い．これは，母語話者ではアクセント指令は文節中に高々 1 個か 2 個であるのに対し，中国語話者はより多くの指令が生起することを示す．図 8 から，中国語話者では最大で 4，5 回まで文節中にアクセント指令が存在する．図中，'0' は当該文節にアクセント指令が存在しないことを示す．これは，文節が 3 モーラ以内の短い文節であるか，アクセント指令の振幅が極端に抑制されている場合である．'0.5' は隣接する文節と

表 5 文発話に生起する中国語話者，母語話者，男女毎のアクセント指令の数の平均値と標準偏差．各 5 話者 × 7 文発話．

	CH_male	CH_female	JP_male	JP_female
mean	9.69	10.71	6.54	7.83
SD	2.46	3.28	1.48	1.81

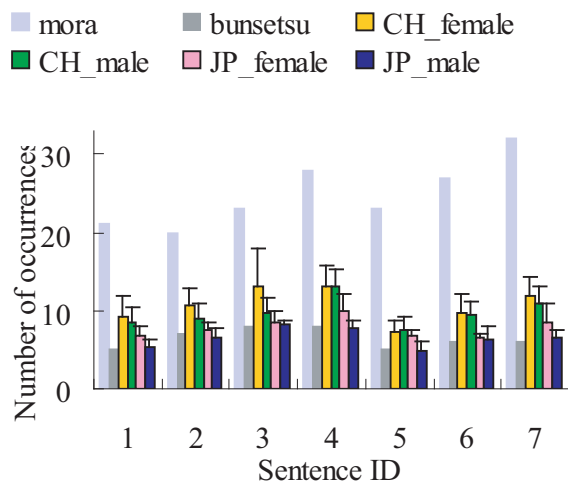


図 7 7 文発話に出現するアクセント指令数の平均値．4 グループ (中国語話者女性・男性，母語話者女性・男性) 各 5 名による発話．棒グラフ上のバーは標準偏差．グラフ左側の 'mora' と 'bunsetsu' は各文中のモーラ数，文節数．

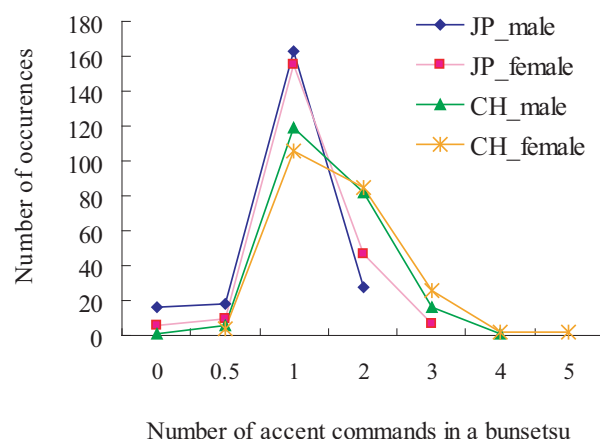


図 8 各文節内に出現するアクセント指令数

アクセント指令を共有していることを意味する．中国語話者にはこのようなケースが少ない．

4.6 負のアクセント指令

日本語の F_0 モデルは，日本語話者の音声の F_0 パターンをきわめて良く近似することが知られているが，背景に様々な母語をもつ日本語学習者の音声の F_0 パターンを十分に良く近似するという保証はない．実際，モデルのパラメータ抽出に当たり，曲線へのフィッティングがうまくいかない箇所が要所に存在した．具体的には，中国語話者の多くの発話において，/eN/，/ai/ など，重音節で終わる文 (節) 末の音節にはアクセントが置か

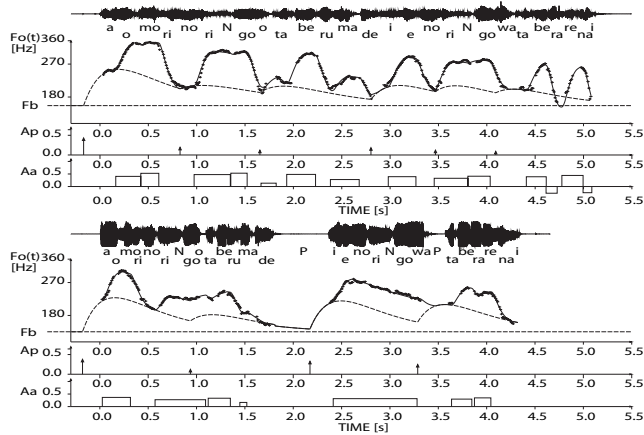


図9 中国語話者(上)と母語話者(下)各1名の発話による F_0 パターンの解析。基底周波数(F_b)、フレーズ指令(A_p)、アクセント指令(A_a)の抽出。発話は「青森の/りんごを/食べるまで/家の/りんごは/食べられない」の7文節から成る。

れやすく、その場合、当該音節のみで完了する F_0 の急激な下降が見られる。これは母語である中国語の声調(トーン)の第4声(下降調)の干渉が考えられる。そこで、標準中国語の基本周波数パターン生成過程モデル[12]に用いられる負のアクセント指令を一時的に導入し、フィッティングを行った。そのような処理が必要であった箇所は、70文の発話中、48箇所存在した。

4.7 考 察

中国語話者は母語話者より、基底周波数が高く、フレーズ指令・アクセント指令とも冗長的に数が多いことが明らかとなった。また、正のアクセント指令に加え、局所的に負の指令も必要であった。図9に中国語話者と母語話者の女性1名ずつの1文発話の例を示す。

発話は6文節から構成されており、中国語話者では文節境界毎にフレーズ指令が生起している。故に表層の F_0 パターンで6つの韻律句を形成する。母語話者は4つのフレーズ指令が生起し、いくつかの文節にまたがって4つの韻律句を形成する。

アクセント指令は、母語話者では7つの正の指令が存在し、通常1文節に1つのアクセント指令が生起することが多いが、ここでは「食べる-まで」と「食べ-られない」の内容語-機能語にそれぞれアクセント指令が与えられている。ただし、これに付与されるアクセント指令は、後続の指令の振幅が小さく抑制されているか、互いが近接している。一方、中国語話者は、文節毎にほぼ2つ以上のアクセント指令を持ち、形態素境界や音節境界毎に指令が生起している。「食べる-まで」と「食べ-られない」の内容語-機能語にそれぞれ付与されたアクセント指令は後続の振幅が抑制されていず、互いに近接もしていない。さらに、負の指令も生起している。また「家の/りんごは」は、母語話者では2文節に1つのアクセント指令が生成されているが、中国語話者では独立した指令が与えられている。

母語話者のある一定時間の間隔を空けたフレーズ指令の生成は、表層の F_0 パターンに大きな韻律的なまとまり(韻律句)を形成させる働きを持つ。頻繁なフレーズ指令の生成は、音声に豊かな抑揚をもたらしうるが、必要以上に頻度が高く、画一的

に生成されると不自然な聴覚印象を与える。また、母語話者のアクセント指令が、1文節に1つ程度与えられる、あるいは後続の指令が小さく抑制されることは、上昇からなだらかな下降に向かう、局所的な丘状の F_0 パターンの形成に役立っていると考えられる。中国語話者の、文節中に比較的大きな指令が並ぶことや、負の指令を持つことは、 F_0 の急激な変化パターンを形成させ、日本語としての発話の自然性を損ねやすい。

5. まとめと今後の課題

中国語話者の日本語発話における韻律的特徴を、 F_0 の観点から、表層のパターンについて概観を述べ、生成過程に注目した分析結果について報告した。今回、基底周波数とフレーズ指令・アクセント指令の生起頻度について結果を述べたが、指令の大きさ、立ち上がり・立下りの時間的タイミングについて分析を進めると、表層の F_0 のレンジ変化パターンとの対応関係が報告できるであろう。被験者数を増やす必要もあり、次には異なる母語を背景とする学習者の音声を F_0 モデル上で比較する予定である。

文 献

- [1] 独立行政法人日本学生支援機構, “留学生受入れの概況(平成17年版)” http://www.jasso.go.jp/statistics/intl_student/data05.htm (accessed 15 Oct, 2006)
- [2] 日本語教育学会(編), “日本語教育機関におけるコースデザイン”, 凡人社, pp.29-37, 1991
- [3] 猪塚恵美子, “学習者に発音を意識させましょう!” 日本語教育研究所主催講演, May 2005.
- [4] 谷口聡人, “音声教育の現状と問題点—アンケート調査の結果について—”, 日本語音声の韻律的特徴と日本語教育—シンポジウム報告—, 重点領域研究「日本語音声」D1 班平成3年度報告書, 1991
- [5] Hirano, H. and Kawai, G., 2005. Pitch patterns of intonational phrases and intonational phrase groups in native and non-native speech, Proc. Interspeech2005, Lisbon, Portugal, pp.761-764.
- [6] 平野宏子, 広瀬啓吉, 峯松信明, 河合剛, 中国語話者の日本語朗読音声の韻律的特徴と母語話者評価, 電子情報通信学会音声研究会 SP2005-188, pp.23-28, 2006.
- [7] Hirano, H., Gu, W., and Hirose, K., Model-based Analysis of F0 Contours of Japanese Sentences Uttered by Chinese Speakers The 7th Phonetic Conference of China (7th PCC) and International Forum on Phonetic Frontiers, to appear, 2006.
- [8] Boersma, P. and Weenik, D., 2006 Praat: doing phonetics by computer <http://www.praat.org/> (accessed 15 Oct, 2006).
- [9] Fujisaki, H. and Hirose, K., Analysis of voice fundamental frequency contours for declarative sentences of Japanese, J. Acous. Soc. Japan (E) 5 (4), pp. 233-242, 1984.
- [10] 成澤修一, 峯松信明, 広瀬啓吉, 藤崎博也, 音声の基本周波数パターン生成過程モデルのパラメータ自動抽出法, 情報処理学会論文誌, Vol.43, No.7, pp.2155-2168, 2002.
- [11] K.Hirose, H.Fujisaki, and S.Seto, A scheme for pitch extraction of speech using autocorrelation function with frame length proportional to the time lag, Proc. IEEE ICASSP, Vol.1, pp.149-152, 1992.
- [12] Fujisaki, H., Hirose, K., Hallé, and P., Lei, H., Analysis and modeling of tonal features in polysyllabic words and sentences of the Standard Chinese, Proc. ICSLP 1990, Kobe, Japan, pp.841-844, 1990.