

生成過程モデルに基づく母語話者と中国語話者の日本語朗読音声の基本周波数パターンの比較*

平野宏子, 広瀬 啓吉 (東大), 河合 剛 (北大), 顧 文涛 (香港中文大), 峯松 信明 (東大)

1 はじめに

国内で日本語を第二言語として学ぶ中国語話者は全体の3分の2以上を占める。我々は、これまで母語話者と中国語話者の日本語朗読音声に現れる韻律的特徴を、句毎のF0パターンの形状や文中のバリエーション、レンジの変化から観察した。本研究では、基本周波数パターン生成過程モデルを用いて、指令の生起頻度や大きさといった観点から特徴を分析する。

2 音声資料

中国語話者・母語話者、男女各15名、合計60名が同一の日本語文7文を朗読した(1文平均25モーラ)。尚、中国語話者・母語話者男女各5名計20名がデータセット1に、残りがデータセット2に含まれる音声で、異なる時期に収録された。全体で、中国語話者は、日本語学習中上級レベル、在日期間2ヶ月 - 30ヶ月の標準中国語話者であり、母語話者は、東京方言話者である。原稿は、ルビ付きの漢字かな混じり文で、読点を含まず、話者が自由に発話した。言い淀みや言い間違いは監督者の指示により、または本人が満足するまで、文頭から読み直した。

3 基本周波数パターン生成過程モデル

基本周波数パターン生成過程モデルは、対数軸上で表現したF0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する[1]。フレーズ成分は、句頭から句末にかけて緩やかに下降する成分に対応し、インパルス状のフレーズ指令の応答として、アクセント成分は、個々の単語やその連鎖に付随する局所的な起伏に対応し、ステップ状のアクセント指令の応答として表現される。

3.1 パラメータ抽出

音声収録(データセット1: 48kHz, 16bit, データセット2: 16kHz, 16bit)後、両セットを10kHz, 16bitにダウンサンプリングし、モデルのパラメータを自動抽出[2]したのち、音素ラベリング、文節境界といった言語情報、波形、F0パターンを参照しながら、手修正を行った。

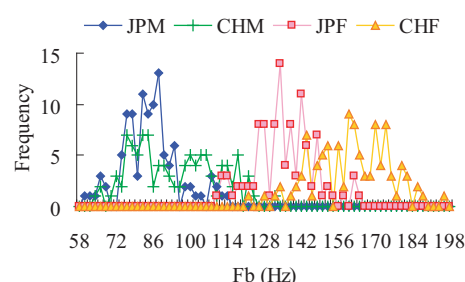


Fig. 1 基底周波数のヒストグラム。JPM/JPF: 母語話者男性/女性, CHM/CHF: 中国語話者男性/女性。

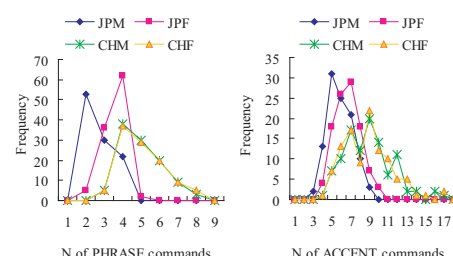


Fig. 2 1文中のフレーズ指令(左)とアクセント指令(右)の生起数。

3.2 基底周波数

中国語話者と母語話者の男女別の各105発声から、基底周波数を求め、ヒストグラムを作成した(図1)。男性の平均値は、中国語話者92.96Hz(SD 15.77), 母語話者84.18Hz(SD 10.70), 女性の平均値は、中国語話者159.80Hz(SD 14.33), 母語話者134.38Hz(SD 11.65)であり、中国語話者の方が母語話者より基底周波数が高い傾向にあることが分かった。1%有意水準のt検定の結果、男女ともに言語間で有意な差が示された。

3.3 フレーズ指令・アクセント指令の生起数

中国語話者と母語話者の各発声から、文毎の各指令の生起数を算出した(図2)。結果、どちらの指令においても、中国語話者の方が多く生起していることが分かった。1%有意水準のt検定では、中国語話者の男女間には違いが見られず、母語話者の男女間に有意な違いが見られた。

文別の生起数(図3)からは、フレーズ成分は、母語話者では文節数の半分程度であるが、中国語話者は、場合によっては全ての文節境界毎に生起しているの

* Model-based analysis of F0 contours of Japanese sentences read aloud by Japanese and Chinese speakers. by HIRANO, Hiroko, HIROSE, Keikichi (University of Tokyo), KAWAI, Goh (Hokkaido University), GU, Wentao (Chinese University of Hong Kong), MINEMATSU, Nobuaki (University of Tokyo)

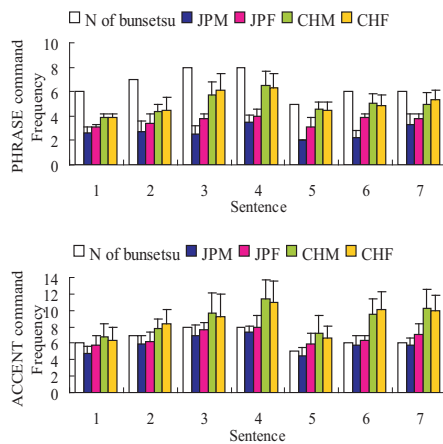


Fig. 3 文毎のフレーズ指令 (上) とアクセント指令 (下) の生起数．白グラフは各文の文節数．

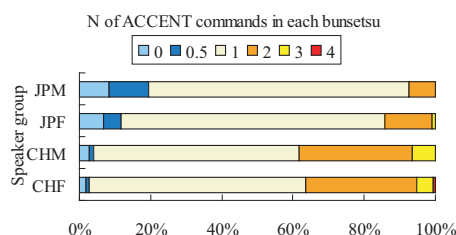


Fig. 4 文節中のアクセント指令の生起数と割合．

が分かる．アクセント指令は，母語話者は文節数と同程度であるが，中国語話者は，文節数よりも多い．文節内のアクセント指令の生起数 (図 4) を見ると，両言語話者とも 1 が基本であるが，中国語話者は 2 以上の割合も高く，逆に，1 アクセント指令を 2 文節が共有する 0.5 (平板式 + 起伏/平板式の 2 連) や，当該文節に指令を含まない 0 (起伏式 + 起伏/平板式で後者の F0 が小さく抑制される場合) は少ない．

3.4 フレーズ指令・アクセント指令の大きさ

各文の文節毎の指令の大きさの変化 (各話者の平均値) を 2 例，図 5 に示す．母語話者は，文節毎に大きさの変化が大きいことがグラフのジグザグから読み取れる．中国語話者は文節毎に変化が少なく，比較的平らなグラフになっている．中国語話者の男女間で違いは見られないが，母語話者では，女性の方が男性よりフレーズ指令が文頭大きく始まるが，その分アクセント指令は男性の方が大きくなるといった傾向が見られた．

3.5 負のアクセント指令の導入

中国語話者の F0 パターンには，特に文 (節) 末の重音節 /-tai/， /-nai/， /-sen/ などに，比較的短時間で急峻な下降幅の大きい F0 パターンがよく出現した．通常の日本語モデルの正のアクセント指令のみでは良好な近似が不可能な場合が多い．中国語のモデル

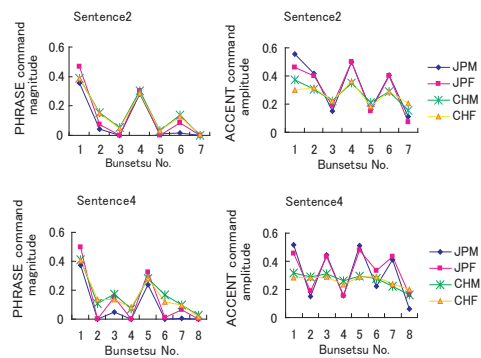


Fig. 5 各文 (上下 2 文) の文節毎のフレーズ指令 (左) とアクセント指令 (右) の大きさの変化．

に想定されている負の指令を導入してフィッティングを行ったところ，良く近似された．このような箇所は，210 発話中 95 あった．

4 まとめと今後の課題

基本周波数パターン生成過程モデルを用いて標準中国語話者と東京方言話者の日本語朗読音声进行分析した．男女各 15 名，計 60 名の 7 文発話の分析から，1) 中国語話者は男女とも，母語話者より基底周波数が高く，2) フレーズ指令，アクセント指令がどちらも過剰生成され，3) 指令の大きさが文節毎に非調整的であり，4) 局所的に負のアクセント指令を必要とすることが示された．また，5) 母語話者男女の指令の生成数，大きさの変化に違いが見られた．1)-3) は非母語話者性を知覚させ，2)-3) は聞き手側に韻律的なまとまりの再構築を要求し，4) は話者が意図しないパラ言語的な情報が付加される可能性がある．

今後，中国語話者の特徴が，母語の干渉であるか，他言語話者にも共通して生じる学習過程の特徴であるかを，母語の韻律，背景母語が異なる話者の特徴を共通のモデル上で調べるにより明確にする．

謝辞 本研究は日本学術振興会の特別研究員奨励費 (194087) の助成を得た。

参考文献

- [1] Fujisaki, H. and Hirose, K., "Analysis of voice fundamental frequency contours for declarative sentences of Japanese", J. Acous. Soc. Japan (E) 5 (4), 233-242, 1984.
- [2] Narusawa, S., Minematsu, N., Hirose, K., and Fujisaki, H., "A method for automatic extraction of model parameters from fundamental frequency contours of speech", Proc. ICASSP, Orlando, FL, vol.1, 509-512, 2002.