

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 30 日現在

機関番号：62618

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23520483

研究課題名(和文) 自発音声データの定量的解析による日本語韻律構造理論の再構築

研究課題名(英文) Reconstruction of the theory of Japanese prosodic structure based upon the quantitative analysis of spontaneous speech data

研究代表者

前川 喜久雄 (Maekawa, Kikuo)

大学共同利用機関法人人間文化研究機構国立国語研究所・言語資源研究系・教授

研究者番号：20173693

交付決定額(研究期間全体)：(直接経費) 3,400,000円、(間接経費) 1,020,000円

研究成果の概要(和文)：非流暢性要素については、filled pauseの高さが、前後の言語的toneから予測可能であることを明らかにした。Filled pauseにはtoneを指定する必要がなく、通常の韻律構造の一部に含める必要はないことが明らかになった。自然下降については、final loweringを分析し、その生起領域が通常は発話を構成する最後のアクセント句であること、ただし、発話の深い切れ目では、もうひとつ前のアクセント句まで領域が広がることを明らかにした。これによって1980年代から続く論争を解決に導いた。副次的成果として韻律特徴の生起頻度情報から話者の性別と年代を精度よく予測できることを示した。

研究成果の概要(英文)：As for disfluency, it was shown that the height of filled pauses can be predicted given the linguistic tones of the preceding and following accentual phrases. Prosodic structure of utterance needs not to include the specification of tone for filled pauses. As for declination, the domain of final lowering was examined. It turned out that the domain of final lowering was the final accentual phrase of an utterance, but in the utterances with deep syntactic boundary, the domain could include the penultimate accentual phrase. This finding put an end to long-standing debate about the domain of final lowering. Lastly, statistical analyses of the occurrence frequency of prosodic events revealed that it was possible to predict the gender (sex) and age of the speakers with the accuracy of about 70% (in the case of gender) and 60% (in the case of age group).

研究分野：音声学

科研費の分科・細目：言語学

キーワード：韻律構造 日本語 filled pause final lowering

1. 研究開始当初の背景

日本語音声の韻律特徴の研究では、言語情報の研究が優先されてきた。そのために、パラ言語情報と関わる要素のなかには、十分な検討をうけてこなかった要素が残されている。

例えば自然な発話のなかには多数の言い淀みや言い直し等の非流暢性要素が含まれているが、従来の研究では、これらの要素はほとんど全く分析されていない。これは実験環境でデータを収集することで、この種の非流暢性要素を積極的に排除してきたからであるが、その結果、現在の韻律構造理論は、そのなかで非流暢性要素をどう位置づければよいかについて何も情報を与えないものになってしまっている。

2. 研究の目的

この研究は、非流暢性要素や、発話の進行に伴うピッチ下降などの現象をとりあげて、日本語の韻律構造と関係づけながら分析することを試みる。大規模な自発音声コーパスである『日本語話し言葉コーパス』のコア部分(CSJ-Core)に記録されたF0情報の定量的解析によって、従来、朗読音声の分析に基づいて構築されてきた発話の韻律構造およびイントネーションについての理論を再検討し、自発音声の諸現象を説明しうる新しい韻律構造を提案する。

3. 研究の方法

特に非流暢性要素(disfluency)と音声基本周波数の自然下降(F0 declination)に関する問題に焦点をあてて検討を進める。

いずれの場合も、CSJ-Coreに付与された韻律アノテーションであるX-JToBIデータを検索して必要な情報を入手する。Filled pauseの場合、CSJ-Coreには10万個以上のfilled pauseが記録されているので、それを分類するとともに、各filled pauseの中央部分のピッチ情報を抽出した。さらに当該filled pauseの直前に位置する通常のアクセント句末尾のtoneのピッチ情報と、当該filled pause直後の通常のアクセント句冒頭のtoneのピッチ情報も抽出した。これら3個のピッチ情報間にどのような関係が成立しているかを探ることが研究の主眼である。

自然下降については、自然下降を構成するいくつかの要素のなかからfinal loweringに注目することにした。Final loweringは発話末尾で、発話の終了を示すために、ピッチが局所的に下降して、話者のピッチ・レンジの下限に達するとされている現象である。

Final loweringは従来の韻律構造理論においても考慮されていることの多い現象だが、その生起領域については、それを発話の最終モーラに限定する節と、より広い範囲が領域になるとする説とが1984年以来、対立したままとなっているので、自発音声データを用いて、その生起領域を検討することにした。

また一部の研究者はfinal loweringは朗読音声に固有の特徴であり、自発音声には存在しないと述べているので、この主張の妥当性もあわせて検討することにした。

自発音声を利用してfinal loweringを研究する際に問題になるのが、発話の定義である。自発音声には句読点などの外的指標は存在しないので、発話の境界を、研究者自身が定義する必要がある。CSJ-Coreには日本語の形態論上の特徴に注目した節境界ラベルが付与されているので、それを利用して、発話境界を、絶対境界・強境界・弱境界の3段階に分類した。

4. 研究成果

非流暢性要素については、最初にfilled pauseの高さ(ピッチ)の平均値が、filled pauseの生起位置とどのように相関しているかを分析した。その結果、2個のアクセント句に挟まれて存在するfilled pauseの高さは、前後いずれのアクセント句に近いかによって、組織的に変動している可能性がうかがわれた。

そこで、前後のアクセント句のtone情報から当該filled pauseの高さを予測できるかどうかを定量的に検討した。比較的シンプルな予測モデルを5種類作成し、それぞれによる予測精度(平均予測誤差)を比較した結果、前後のtone間を直線補完するモデルが最も予測精度が高いことが判明した。また予測誤差の絶対値がfilled pauseの平均的なピッチレンジの幅以内に収まる水準のものであることも明らかになった。

これはfilled pauseの高さが、前後の言語情報から予測可能であり、従って、韻律構造上は指定する必要がない(underspecifyできる)ことを示唆する結果である。Filled pauseの高さの指定に関する限り、従来の韻律構造には大きな変更は必要ないと考えられる。

次にfinal loweringについて述べる。ここでは、いわゆるdownstepに起因するデータの複雑化を避けて、データを斉一的に検討可能とするために、すべて有核のアクセント句のみから構成される発話だけを検討することにした。発話の長さについては、アクセント句数にして1個から5個までの発話を検討した。これ以上長い発話はデータ数僅少のために検討することができなかった。また、すべて無核アクセント句からなる発話も検索したが、やはりデータ数僅少のために検討を断念した。

話者ごとにZ変換された対数F0空間上で、有核アクセント句を構成する4個のtone(発話冒頭の低toneであるILT、発話冒頭のピッチ上昇の頂点であるIHT、アクセント核の高さを示すAcc、そして発話末の低toneであるFLTの4個)をプロットすることで以下の事実が明らかになった。

(1) アクセント句数2個から5個から構成さ

れる発話では、N-1 個目のアクセント句から N 個目の（つまり最終）アクセント句にかけて顕著なピッチの下降が生じている。

- (2) その際、最終アクセント句を構成する 4 個の tone はすべて一律に下降している。
- (3) 発話の長さに関わらず、最終アクセント句はほぼ同一の高さに位置している。

これらの事実は、自発音声にも final lowering が明瞭に観察されることを示すとともに、final lowering の生起領域が発話の最終アクセント句全体であることを明瞭に示している。ただしアクセント句 1 個のみから構成される発話はやや例外的な振る舞いを示していた。この種の発話には final lowering が十分に現れず、発話末のピッチは話者のピッチレンジの下限よりもかなり高めに位置していた。これはアクセント句 1 個からなる発話では、発話の冒頭アクセント句が同時に最終アクセント句であるという特殊事情によって生じた例外的ふるまいであると解釈することができ、上記の結論に抵触するものではないと考えられる。

次に発話の切り出しに利用した節境界の深さと final lowering の関係も検討した。Final lowering は 3 種類の発話のいずれにおいても観察されることが確認でき、final lowering が典型的な文末だけでなく、より広範な環境に生じていることが判明した。

ただし発話の最終モーラに生じる FLT の高さに注目すると、絶対境界において最も低く、強境界がそれにつき、弱境界において最も高い値をとっているため、final lowering の程度は統語的な境界の強さによって変化しているものと考えられた。

さらに興味深い発見は、絶対境界発話においては、最終アクセント句だけでなく、次末（最後から 2 番目）のアクセント句にまで final lowering の効果が及ぶことが分析によって確認されたことである。つまり、final lowering は統語的環境に応じて、音声学的な強度を変化させると同時に、その生起領域も変化させているのではないかと考えられる。

今回の研究は日本語（東京方言）の final lowering について過去 20 年間未解決であった生起領域の問題に大きな光を投げかけるものとなった。

最後に、副次的成果について述べる。CSJ-Core の X-JToBI アノテーションには多岐にわたる韻律現象が記録されている。それらの現象の生起頻度を講演ごとに集計した頻度データを作成して、判別分析を適用すると、話者の性別および年齢をかなりの精度で予測できることが判明した。性別であれば 70% 以上、年齢（10 歳刻みにまとめた年代）であれば、60% 前後の精度が得られる。精度の高さはともかく、記号化された韻律情報だけを利用することで、チャンスレベルをはるかに上回る予測が可能であることから、発話の韻律情報には、現在我々が想定しているよりも

さらに多くのパラ言語情報が含まれている可能性を示唆する結果として解釈できる。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 2 件)

K. Maekawa. "Prediction of Non-Linguistic Information of Spontaneous Speech from the Prosodic Annotation: Evaluation of the X-JToBI System." *Proc. LREC 2012*, pp.991-996, 2012:05:23.

K. Maekawa. "Prediction of F0 height of filled pauses in spontaneous Japanese: A preliminary study." *Proc. DiSS 2013 (The 6th Workshop on Disfluency in Spontaneous Speech, Stockholm*, pp.41-44, 2013:08:25.

〔学会発表〕(計 2 件)

K. Maekawa. "Domain of final lowering in spontaneous Japanese." *J. Acoustical Society of America*, 135(4), p.2194, 2014:05:05.

前川喜久雄「自発音声中のフィラーの特性に関する予備的分析:位置と高さの分析」第 26 回日本音声学全国大会予稿集, pp.115-120, 2012:09:29.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況 (計 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

http://researchmap.jp/MAEKAWA_Kikuo/

6 . 研究組織

(1)研究代表者

前川 喜久雄 (MAEKAWA, Kikuo)

研究者番号 : 20173693

(2)研究分担者

()

研究者番号 :

(3)連携研究者

()

研究者番号 :