

日本語自発音声における final lowering の生起領域

前川 喜久雄 (国立国語研究所言語資源研究系)
kikuo@ninjal.ac.jp

1. 問題のありか

日本語の東京方言には final lowering が存在するといわれている。Final lowering (以下 FL)とは、発話の終了を示すために発話末尾で F0 が話者のピッチレンジの下限まで下降するとされている現象である。FL に関する科学的な研究は英語イントネーションモデルの一部として FL を検討した Liberman and Pierrehumbert (1984)を嚆矢とする。日本語に関しては Poser(1984)、Pierrehumbert and Beckman (1988)などの研究がある。これとは独立に日本語の音声学では、古くから平叙文末尾のイントネーションが下降調であるか平調であるかをめぐっての議論がある (郡 2008, 2010 参照)。これも FL と関連する問題である。

Pierrehumbert and Beckman (以下 P&B)は、日本語イントネーションにおける F0 下降の主要な言語的要因として、時間の関数である declination (いわゆる自然下降)、アクセント核によるピッチレンジの狭めである catathesis (downstep)とならんで、FL を挙げている。しかし、日本語イントネーションのモデルがすべて FL を想定しているわけではない。藤崎による対数 F0 重畳モデルには FL に対応するコンポーネントが存在しないし(Fujisaki and Hirose, 1984)、NTT 基礎研究所で開発されたモデルも同様である (佐藤,1987)。日本語に FL が存在するかどうかは改めて検討する価値のある問題である。

また日本語に FL を認めるモデルのなかでも、FL の作用域についての見解に対立がある。Poser は FL の作用域を発話の最終モーラとしているが、P&B は FL がより広い領域に生じている可能性を少数の実例に基づいて示唆している。いずれを採用するかはイントネーションモデルの設計に大きく影響する。

ところで上述の先行研究はすべて朗読音声进行分析したものであった。これに対して Umeda(1982)はイントネーションの F0 下降にはスタイル依存性があるとして、FL などが朗読音声に固有の現象である可能性を示唆している。そもそも自発音声に FL が生じるかどうか、これもまた経験的な検討に値する問題である。これらの問題について筆者は先に『日本語話し言葉コーパス』のコア (以下 CSJ-Core) を利用した検討結果を報告した (Maekawa, 2010)。今回、同じ CSJ-Core を用いてより一般的な結果を得たのでここに報告する。

2. データ

CSJ ないし CSJ-Core と X-JToBI によるアノテーションについては既に多くの場で紹介してきているので省略に従う (詳細は前川 2004、五十嵐ほか 2006 など参照)。CSJ のアノテーションのうち今回の分析で重要な役割を果たすのは節境界ラベル(CBL)である。CSJ の節単位情報アノテーションでは、節を 49 種に分類したうえで、統語的な強さに従って3つのクラスに分類している。すなわち絶対境界 (「～デス、～マシタ」等のいわゆる文末表現)、

強境界（～ガ、～ケレドモなど、大略、南不二男のC類従属節）、弱境界（～タラ、～レバ、～カラ、～ノデ、～デ、～トイウなど）である。詳しくは丸山ほか(2006)を参照のこと。

また、今回の分析には最近公開を開始したCSJ-RDB版を利用した。この版と従来筆者らが研究に利用してきたX-JToBIデータ(菊池・前川2007)とでは、アクセント句(AP)の認定に組織的な相違がある。例えば従来のデータではフィラー(filled pause)は隣接するテキストと一体化して1個のAPを構成していたが、CSJ-RDB版では隣接するテキストから分離されて単独で韻律上の単位を構成しているなどの点である。詳しくは小磯ほか(2013)参照。旧データと新データでFLの分析結果に大きな異同が生じるとは考えにくい、この点を確認することも本研究の副次的な目的である。

以下の分析では、CSJ-Coreのうち自発音声である学会講演(APS: 18.7h, 77864AP, 男性話者46名, 女性話者24名)、模擬講演(SPS: 19.9h, 80847AP, 男性53名, 女性54名)、対話(3.7h, 9236AP, 男性9名, 女性9名)を対象とした。再朗読データは利用していない。

3. 分析

3.1. 発話のきりだし方

自発音声において何を発話として認定するかは研究者によっても、またデータによっても異なりうる。今回は上述の節境界ラベルを利用して数種類の異なる「発話」をデータからきりだした。両端を絶対境界で区切られ、内部に絶対境界ラベルを含まないという条件で切り出されるアクセント句の連鎖を以下では絶対区分発話と呼ぶ。同様に強境界ラベルで区切られた発話（内部に絶対および強境界ラベルなし）を強区分発話、弱境界ラベルで区切られた発話（内部に節境界ラベルなし）を弱区分発話と呼ぶことにする。

3.2. 有核句連鎖の分析

日本語のAPには有核句が圧倒的に多いので(7割以上)、まず分析対象を有核句のみから構成されている発話に限定した分析を行う。有核句と無核句が混在した発話ではcatathesisの効果の評価が複雑化するが、有核句連鎖であればその問題も単純化できる。以下ではX-JToBIで有核句を構成する4個のトーンをILT(Initial Low Tone～X-JToBIの%L)、IHT(Initial High Tone～H-)、Acc(Accent,～A)、FLT(Final Low Tone～L%)の略号で表現する。頭高型の有核句にはIHTは付与されていない。また母音の無声化その他の要因でF0が測定不能なケースが頻発するが、それらのトーンは計算から除外する。

図1-3は1~5個の有核句から構成される絶対区分発話、強区分発話、弱区分発話について、ILT, IHT, Acc, FLTの平均値を示したものである。横軸の記号はトーンの種類に発話を構成する有核句の位置を付加した記号(例えばILT3は発話頭から数えて3番目の有核句のILT)である。縦軸はF0値を話者ごとの平均値と標準偏差を使って正規化(Z変換)した値である。この正規化データを利用するかぎり男女差は実質上消失するので、以下では男女のデータをプールして分析する。図4は比較のために節境界ラベルが一切付与されていないAPの連鎖について、図1-3と同じ分析を施した結果である。

図1-4からは以下の傾向が読みとれる。①いずれの区分発話タイプ(絶対、強、弱)にお

いても、そしてどの発話長（AP 数 1~5）においても、発話末尾の FLT はほぼ同一の値まで

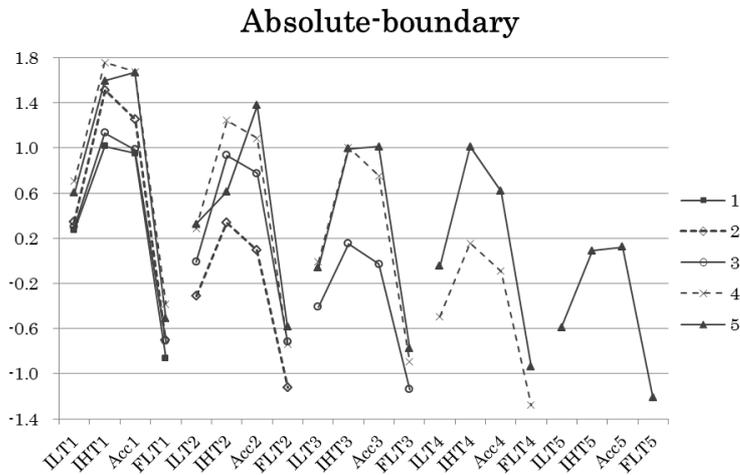


図 1. 絶対区分発話有核句連鎖のトーン平均値(正規化 F0)

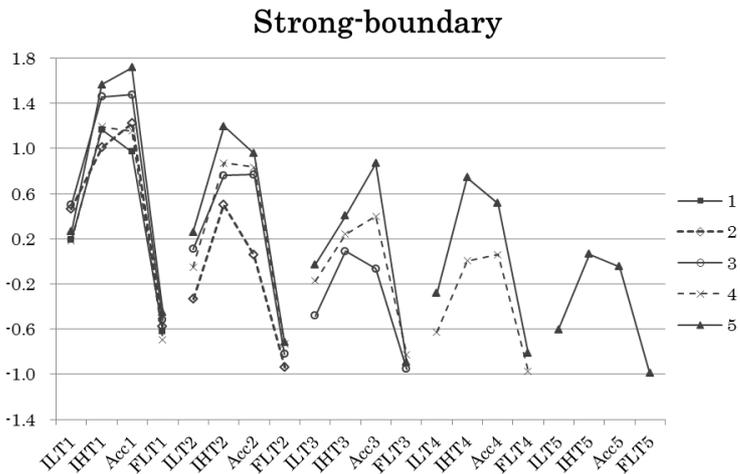


図 2. 強区分発話有核句連鎖のトーン平均値(正規化 F0)

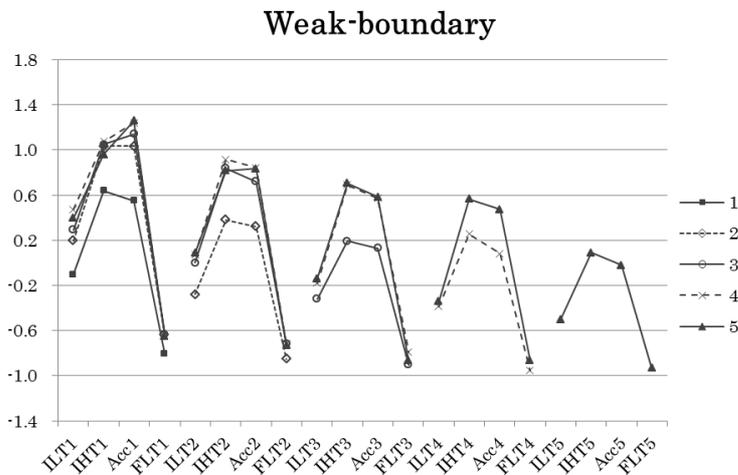


図 3. 弱区分発話有核句連鎖のトーン平均値(正規化 F0)

下降している。②その下降は主に発話の最終 AP が全体として下降することによって生じており、最終 FLT だけが下降しているのではない。③ただし AP 1 個のみからなる発話では AP 全体も FLT も他の発話に比べると十分に下降しない。④また節境界が存在しない環境 (図 4) では末尾 AP に特異的な下降 (FL) そのものが生じない。⑤そして最終 AP の FLT 値は発話タイプによって変動しており ABS<STR <WEAK の関係にある。一部に無核句を含む発話で試行しても同じ結論が得られた。このうち①と②は Maekawa(2010)の結果と一致しており、③以降は今回の新知見である。

3.3. 節境界の効果

新知見のうち④,⑤についてももう少し検討する。まず FL によって下降するのは最終 AP だけかながえてよいかを検討する。図 5 は 3 種類の節境界および節境界無し (None) の AP で構成される区分発話について、節末に位置する AP を構成するトーンの正規化 F0 平均値 (ILT0~FLT0)、節末よりひとつ前に位置する AP (次末 AP) を構成するトーンの平均値 (ILT-1

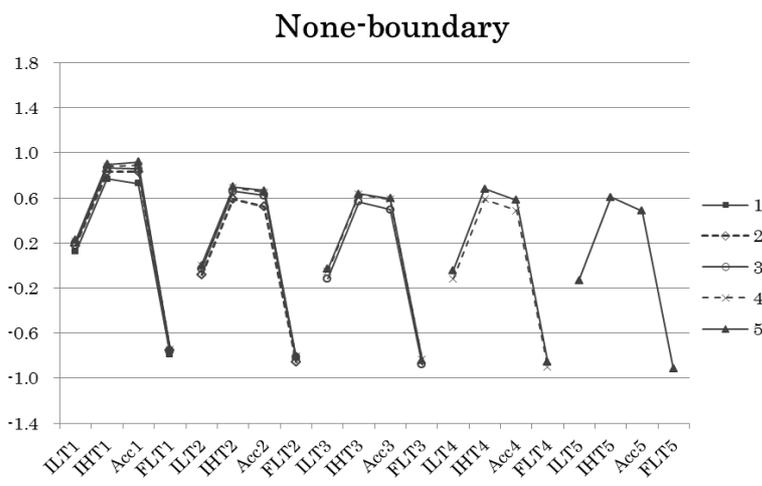


図 4. 節境界が存在しない有核句連鎖のトーン平均値

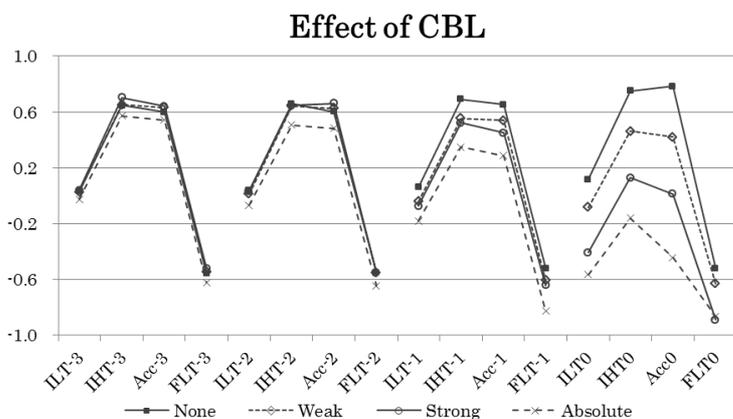


図 5. 節境界ラベルの効果の及ぶ範囲

要因とした 3 要因分散分を行った。その際、全データを利用するとデータ数過多のために (AP に含まれるトーンを別個に扱うのでデータ総数は 64 万個以上になる)、科学上意味のない差にまで有意差が生じてしまうという問題が生じる。そこで検定力分析 (事前分析) を行って最低限必要なサンプル数を推測した。検定力分析の結果を参照して 3000 個のサンプルを無作為抽出し、分散分析に利用した。表 1 に結果を示す。有意性の認められた主効果は節境界ラベルとトーンのみである。ただし AP 位置は主効果としては有意でないが、節境界ラベルとの間の交互作用に有意性が認められた。この交互作用は節境界の効果は最終 AP とそれ以前の AP とで実質的にことなっていることを示していると考えられる。

次いで、節境界ラベルに関して Holm の方法による多重比較を実施した。絶対境界と無ラベル、絶対境界と弱境界の間に 1%水準の有意差が認められたが、絶対境界と強境界を含む他のすべての対には有意差は認められなかった (以上はすべての AP 位置をプールした場合の結果であることに注意)。最後に、AP 位置と節ラベルの交互作用を検討するために、AP

~FLT-1)、二つ前の AP (次々末 AP)を構成するトーンの前平均値(ILT-2~FLT-2)、三つ前の AP (次々々末 AP)を構成するトーンの前平均値(ILT-3~FLT-3)を示したものである。ここで分析しているのは CSJ-Core 全体であり、有核句連鎖に限られていない。この点で図 1-4 とは異なることに注意。

最終 AP における節境界の効果は顕著であるが、次々末 AP にも類似の効果が存在しているように見え、次々末 AP においては節境界が絶対境界であるケースのみ、平均値がやや低下しているように見える。この問題を統計的に検討するためにトーン (ILT, IHT, Acc, FLT の 4 水準)、節境界ラベル (絶対、強、弱、ラベル無の 4 水準)、AP の位置 (0, -1, -2, -3 の 4 水準) を

位置別に節境界の多重比較を行った。各位置ともそれぞれ 3000 個のサンプルを無作為抽出して分析した。結果を表 2 に示す。次末 AP では節境界が絶対境界の場合にだけ下降が認められるが節境界による差が認められるのはここまでである。今回の分析手法によるかぎり、節境界の効果は主に最終 AP に生じ、次末 AP にも部分的に認められるが、次々末 AP 以降には認められない。

表 1. 3 要因分散分析表

	df	Sum Sq	Mean Sq	F value	Pr(>F)
節境界ラベル	3	17.3	5.77	8.357	<0.001
トーン	3	501.6	167.11	242.316	<0.001
AP 位置	3	1.2	0.41	0.597	0.617
節境界ラベル*トーン	9	7.2	0.80	1.156	0.320
節境界ラベル*AP 位置	9	24.9	2.77	4.017	<0.001
トーン*AP 位置	9	7.0	0.78	1.125	0.341
節境界ラベル*AP 位置*トーン	26	10.9	0.42	0.606	0.941
残 差	2051	1415.1	0.69		

表 2. 多重比較の結果

	最終 AP	次末 AP	次々末 AP	次々々末 AP
無境界：弱境界	***	n.s.	n.s.	n.s.
無境界：強境界	***	n.s.	n.s.	n.s.
無境界：絶対境界	***	***	n.s.	n.s.
強境界：弱境界	n.s.	n.s.	n.s.	n.s.
強境界：絶対境界	n.s.	**	n.s.	n.s.
弱境界：絶対境界	**	***	n.s.	n.s.

有意性水準 ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, n.s.: $p \geq 0.05$

4. 結論と今後の課題

第一に今回の分析結果は自発音声にも FL が生起することを明確に示していた。FL は日本語イントネーションの重要な構成要素であると考えられる。第二に CSJ-RDB 版と従来のデータの間には分析結果に関する重要な不一致は見つからなかった。

第三に FL が最終モーラだけを下降させるという主張は正しくないことが明らかになった。FL の生起領域には少なくとも発話の最終 AP 全体が含まれている。ただし、分散分析の結果に示されていたように、AP 位置と節境界タイプの間には交互作用が存在する。すなわち、絶対境界では最終 AP だけでなく次末 AP にも若干の下降が認められるのに対し、他の境界では最終 AP のみが下降している (図 5 と表 2 参照)。P&B が FL の生起領域が最終 AP を超えて広がる可能性を示唆しているのは、絶対文末を観察しているためであると考えられる。

第四に、前項とも関係するが、CSJ の節境界ラベルを利用して種々の発話を規定し、FL との関係を検討した結果、発話最終 AP の下降は絶対境界に限らずすべての節境界で生じていること、そして節境界の深さと相関して下降の程度が変化していることが判明した。こ

これは日本語イントネーション研究全体にとって重要な発見である。日本語の節末では節の統語的な特性を反映した F0 の調整が行われており、従来 FL として語られてきた文末（絶対境界）に生じる FL は、広義の FL のうち最も顕著なものにすぎないことが明らかになったと考えられるからである。

最後に今後の課題に触れる。重要な課題には以下のふたつがあると考えている。第一に FL から declination の影響を差し引いて再評価する方法を考えること。一般に深い節境界ほど発話冒頭からの平均距離が遠くなる傾向が存在するので、今回 FL の効果と見た F0 下降のなかには自然下降による成分も含まれている。それを分離した上で FL を再評価する必要があるのだが、そのためには declination の開始位置（おそらく息継ぎの位置）をどのように決定するかが問題になる。第二に、図 1-4 のような分析において AP 間の修飾関係を考慮した再分析を実施する必要がある。このためには CSJ-Core に付与された文節係り受けアノテーションが利用できると思われる。

謝辞：この研究は文科省科研費基盤研究（課題番号 23520483 および 23320081）および国立国語研究所共同研究「コーパス日本語学の創成」によって実施されたものである。

参考文献

- Fujisaki, H. and K. Hirose (1984). "Analysis of voice fundamental frequency contours for declarative sentences of Japanese." *Journal of the Acoustical Society of Japan (E)*, 5 (4), 233-242.
- Lieberman, M. and J. B. Pierrehumbert (1984). "Intonational invariance under change in pitch range and length." In M. Aronoff and R. Oerhle (eds.) *Language Sound Structure*. MIT Press.
- Maekawa, K. (2010). "Final lowering and boundary pitch movements in spontaneous Japanese." *Proceedings of DiSS-LPSS Joint Workshop 2010, Tokyo*, 47-50.
- Pierrehumbert, J. B. and M. E. Beckman (1988). *Japanese Tone Structure*. MIT Press.
- Poser, W. J. (1984). *The Phonetics and Phonology of Tone and Intonation in Japanese*. Ph.D. Thesis, MIT.
- Umeda, N. (1982). "F0 declination" is situation dependent." *Journal of Phonetics*. 10, 279-290
- 五十嵐陽介, 菊池英明, 前川喜久雄(2006). 「韻律情報」国立国語研究所(2006)所載, 347-453.
- 菊池英明, 前川喜久雄(2007). 「韻律研究のための日本語話し言葉コーパス XML 文書作成」人工知能学会研究会資料 SIG-SLUD-A603, 3-8.
- 小磯花絵, 前川喜久雄, 五十嵐洋介(2013). 「『日本語話し言葉コーパス』における韻律単位の認定基準について」第3回コーパス日本語学ワークショップ予稿集, 351-358.
- 郡史郎(2008). 「東京方言における平叙文末の下降増大現象—平叙文末は平調か下降調か—」音声言語, VI, 79-102.
- 郡史郎(2011). 「東京方言の平叙文文末に見られる下降増大現象について（再考）」音声研究, 15 (3), 76-77.
- 国立国語研究所(2006). 『日本語話し言葉コーパスの構築』国立国語研究所.
- 佐藤大和(1987). 『規則による音声合成の研究』北海道大学学位論文.
- 前川喜久雄 (2004). 「『日本語話し言葉コーパス』の概要」日本語科学, 15, 111-133.
- 丸山岳彦, 高梨克也, 内元清貴(2006). 「節単位情報」国立国語研究所(2006)所載, 255-322.