

句末境界音調のピッチレンジに与える要因： 『日本語話し言葉コーパス』の分析

五十嵐 陽介（広島大学）[†]
小磯 花絵（国立国語研究所理論・構造研究系）[‡]

Factors Affecting Pitch Ranges of Boundary Pitch Movements: An Analysis of the Corpus of Spontaneous Japanese

Yosuke Igarashi (Hiroshima University)
Hanae Koiso (Dept. Linguistic Theory and Structure, NINJAL)

1. はじめに

韻律句 (prosodic phrase) 末尾に生じる音調で発話の語用論的解釈 (質問、継続、強調など) に貢献する音調を句末境界音調 (Boundary Pitch Movement, BPM) という。どのような BPM がいくつあるのかに関する研究や、BPM の機能に関する研究は古くからなされているが (川上 1963, 郡 1997)、BPM のピッチレンジ (pitch range) に関する研究はほとんどなされていない。

日本語の韻律句のピッチレンジを決定する主要な要因のひとつとして、アクセント句のピッチレンジを縮小させるダウンステップ (downstep) と呼ばれる現象が知られている (Pierrehumbert and Beckman 1988)。我々は以前の研究で (五十嵐・小磯 2012, Igarashi and Koiso 2012)、『日本語話し言葉コーパス』 (前川 2004, 以降 CSJ) の分析に基づいて、BPM にダウンステップが観察されるか否かを検討した。その結果、BPM には、アクセント句の主要部 (BPM を除いた部分) に観察されるダウンステップに類似した現象が観察されることが明らかになった。しかしながらこの研究では、話し方のスタイルや韻律句における BPM の位置などの要因が考慮されていなかった。

本研究の目的は、CSJ の分析を通じて、BPM のピッチレンジに効果を与える要因を明らかにすることにある。第 2 節では、BPM の記述のために必要な諸概念を導入したのち、我々の以前の研究 (五十嵐・小磯 2012, Igarashi and Koiso 2012) を中心に、BPM のピッチレンジに関連する研究を再検討する。第 3 節では用いたデータを記述する。第 4 節では分析結果を報告する。第 5 節で結果の考察を行い、結論を述べる。

2. 日本語の韻律構造の記述

2.1 ダウンステップ

日本語におけるダウンステップとは、アクセント核 (lexical pitch accent) が、それを含むアクセント句 (以降 AP) に後続する AP の基本周波数 (F0) 頂点 (ピッチレンジの上限) を、反復的 (iterative) に低下させる現象である (Pierrehumbert and Beckman 1988)。図 1 は「旨い飴がありました」 (左) と「旨い豆がありました」 (右) の音声波形と基本周波数 (F0) 曲線を示したものである。双方の発話とも、最初の AP (ウマイ) はアクセント核を持つ有核句である。そのため 2 番目の AP (アメガ/マメガ) にダウンステップが生じ、F0 頂点が低下する。一方、2 番目の AP は、左の発話ではアクセント核を持たない無核句 (アメガ) であるのに対して、右の発話では有核句 (マメガ) である。したがって、3 番目の AP (アリマシタ) にダウンステップが観察されるのは、右側の発話のみとなる。

[†] igarashi@hiroshima-u.ac.jp, [‡] koiso@ninjal.ac.jp

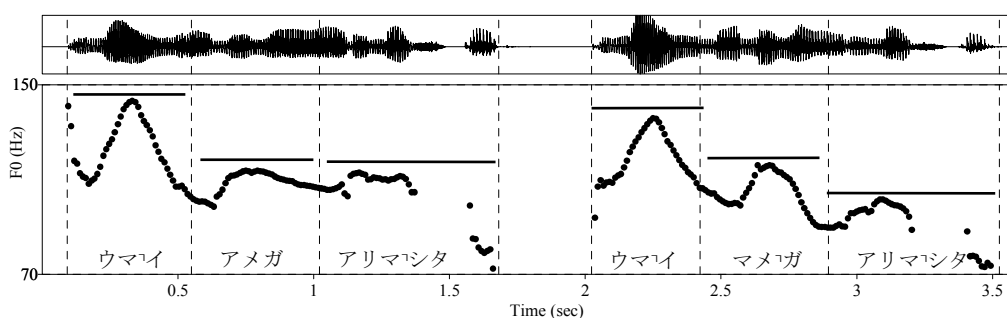


図1 ダウンステップ.

左の発話は「旨い飴がありました」、右の発話は「旨い豆がありました」。縦の点線は AP 境界を表す。各 AP のピッチレンジの上限を水平方向の実線で示している。発話者は第 1 筆者。

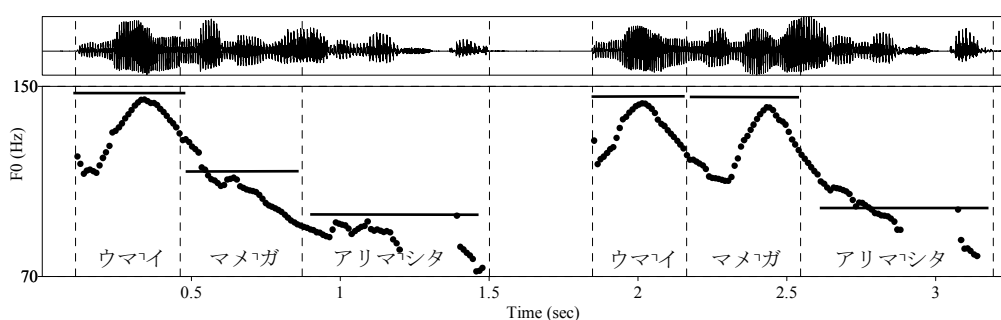


図2 フォーカスによるダウンステップの阻止.

発話は左右とも「旨い飴がありました」。左の発話は「旨い」に、右の発話は「豆」にフォーカスが置かれている。縦の点線は AP 境界。水平方向の実線はピッチレンジの上限。発話者は第 1 筆者。

ダウンステップは、AP の F0 頂点を低下させるが、ピッチレンジの下限はあまり低下させない。このため AP の F0 変化幅は、有核句が先行することにより反復的に縮小することになる。

2.2 イントネーション句

ダウンステップの効果は、一定の統語構造やフォーカスによって阻止されることが知られている (Pierrehumbert and Beckman 1988; Venditti et al. 2008)。図 2 は「旨い豆がありました」の F0 曲線を示したものであるが、左の発話は「旨い」にフォーカスが置かれており、発話は「豆」にフォーカスが置かれている。左側の発話では 2 番目以降の AP にダウンステップが観察される。フォーカスは後続要素の F0 頂点をさらに低下させる効果があるため (post-focal prosodic subordination, cf. Venditti et al. 2008)、図 2 (右) の発話と比較して、F0 頂点の低下の程度がより顕著になっている。

一方図 2 (左) では、フォーカスを受けた語「豆」を含む AP (マメガ) の F0 頂点は、それに先行する AP (ウマイ) の F0 頂点とほぼ同水準となっており、ダウンステップが観察されない。この現象を記述するために、Pierrehumbert and Beckman (1988) の韻律理論では、AP (マメガ) の始端に、AP より階層的に上位の韻律句の境界を仮定する。この韻律句は、CSJ が採用している日本語の韻律ラベリング体系である X-JToBI (Maekawa et al. 2002) およびその前身である J_ToBI (Venditti 2005) では、イントネーション句 (intonation phrase, 以降 IP) と呼ばれおり、ダウンステップの生じる領域、あるいはピッチレンジが指定される領域として定義される。X-JToBI では、図 2 の発話に図 3 に示す韻律階層が仮定される。



図3 図1の発話の韻律階層。

左側は「旨い」にフォーカスが置かれた発話、右側は「豆」にフォーカスが置かれた発話。APより下位の韻律単位は省略してある。

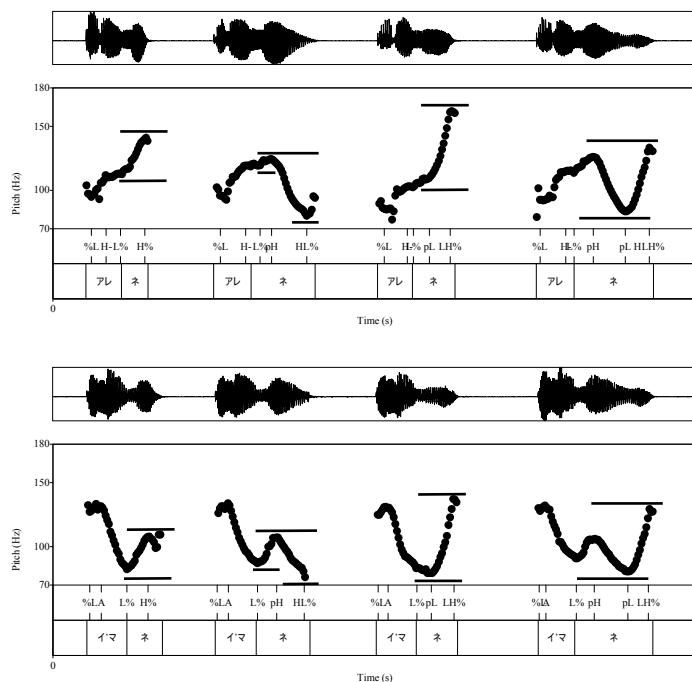


図4 BPM.

上は無核句（アレネ）の句末に BPM を伴う発話（4 発話）、下は有核句（イマネ）の句末に BPM を伴う発話（4 発話）である。BPM の種類は上下ともに左から H%, HL%, LH%, HLH% である。水平方向の実線によってピッチレンジの上限と下限を示している。HL% には下限が発話者は第 1 筆者。

2.3 句末境界音調（BPM）

日本語にどのような BPM がいくつあるかに関する研究は数多くあるが、見解の一致は得られていない（川上 1963, 郡 1997）。日本語の BPM の種類と数を確立することは重要な研究課題であるが、本研究ではこれに取り組まない。本研究では X-JToBI の枠組みに基づいて BPM を記述する。

X-JToBI では主要な BPM として図 4 に示す 4 種類が認められている。H%（上昇調 1）は F0 が単純に上昇するタイプの BPM であり、HL%（上昇下降調）は上昇の後に下降が生じるタイプの BPM である。LH%（上昇調 2）は、F0 が上昇する点は H% と同様であるが、上昇の前に低い F0 が一定時間継続する点が異なる。HLH%（上昇下降上昇調）は、上昇の後下降が生じ、その後さらに上昇が生じるタイプの BPM である（五十嵐他 2006）。

BPM は複数の F0 屈曲点によって特徴づけられるため BPM の物理的実現の分節音列上の生起時刻と F0 値を記述するためには、複数のトーンラベルが必要となる。この目的のために X-JToBI では、BPM のトーンラベルを分解し、複数の屈曲点の生起位置を記述している。この際、トーンラベルを単に分解しただけでは検索に支障をきたすので、分解されたラベ

ルの一部には補助記号を付与することになっている。4 種類の BPM それぞれの記述に用いられるトーンラベルと、そのラベルが記述する物理的なイベントは表 1 に要約されている。トーンラベルを用いた BPM のラベリング例は図 4 に示されている。

表 1 BPM ラベル

BPM	用いられるトーンラベルとそれが記述する物理的イベント
H%	L% (上昇開始点), H% (上昇終了点)
HL%	L% (上昇開始点), pH (下降開始点), HL% (下降終了点)
LH%	L% (低 F0 区間開始点), pL (上昇開始点), LH% (上昇終了点)
HLH%	L% (上昇開始点), pH (下降開始点), pL (上昇開始点), LH% (上昇終了点)

2.4 BPM のピッチレンジ

BPM にダウンステップは観察されるのであろうか。我々はこの問いに答えるために、CSJ Core に含まれる 177 ファイル (約 39 時間) に生じる BPM のピッチレンジを計測した (五十嵐・小磯 2012, Igarashi and Koiso 2012)。その結果、BPM に先行するアクセント核の数が増えるほど BPM の F0 頂点が低下する現象 (ダウンステップ) は、どの種類の BPM にも観察されることが明らかになったが、上昇幅・下降幅の縮小の有無や縮小の様相は BPM の種類ごとに異なることも明らかになった。各 BPM の結果は以下のように要約できる。

H% に関しては、先行アクセント核数の増加にしたがって、上昇の終点である F0 頂点だけでなく上昇の起点も同程度低下するため、上昇幅はほぼ一定に保たれることが分かった。一方 HL% に関しては、BPM の上昇部分と下降部分でピッチレンジ制御が異なることが明らかになった。上昇の起点は、H% の場合と同様に、先行アクセント核数の増加にしたがって反復的に低下するため、上昇幅はほぼ一定になることが明らかになった。一方、下降の終点は、アクセント核数がゼロの場合とそれ以外の場合に差が観察される以外は、ほぼ一定の値に保たれるため、下降幅は、先行アクセント核数の増加にしたがって反復的に縮小することが分かった。LH% には明確な傾向が観察されなかったが、これはこの BPM の総数が少なかったためであると考えられる。(HLH% は十分な数が得られなかったため分析の対象外とした。)

しかしながら、我々はこの研究で、話し方のスタイルや韻律句内での BPM の位置などの要因を考慮しなかった。これらの要因は BPM のピッチレンジ制御に効果を与える可能性がある。前川 (2011) によると、BPM の種類の選択や韻律句内での BPM の位置は、話し方のスタイルと強く関連している。前川 (2011) は CSJ の BPM の出現頻度を分析し、あらたまった発話には H% が、くだけた発話には HL% が相対的に多く出現することを明らかにした。また前川 (2011) は CSJ に含まれる「学会講演」「模擬講演」等の別 (以降これをレジスターと呼ぶ) を分離する変数を検討し、その変数に BPM に関連する特徴が含まれていることを明らかにした。ひとつは BPM の種類である。「学会講演」は H% (および PNLP, EUAP¹) が多いことによって特徴付けられ、「模擬講演」は HL% が多いことによって特徴付けられる。もうひとつは BPM の生じる位置である。BPM は IP の最後に位置する AP 末に生じることが多いが、それ以前の AP の末尾に生じることもある。「学会講演」は IP 末以前に BPM が生じる現象が多いことによって特徴付けられ、「模擬講演」は、そのような現象が少ないことによって特徴付けられる。

前川 (2011) の分析結果は、特定のスタイルに強く結びついたイントネーションパターンが存在していることを示している。そのパターンを規定する特徴に、これまで明らかにされてきた BPM の種類や、BPM の出現する位置だけでなく、ピッチレンジを含めた BPM の音声実現の違いが含まれている可能性を検討してみることは有益であるように思われる。

¹ PNLP と EUAP については五十嵐他 (2006) 参照。

3. データと計測

データとして使用したのは CSJ Core に含まれる「学会講演」70 ファイル (19 時間) と「模擬講演」107 ファイル (20 時間) である。以下、用いたデータと計測方法について詳述する。

3.1 分析対象

4 種類の BPM のうち LH% (286 件) と HLH% (7 件) は数が少ないので除外した。先行アクセント核数 (3.2 参照) が 6 以上の BPM は除外した。AP の次末モーラ以前から上昇を開始する BPM の変種も対象から除外した。H% には、F0 上昇終了後に同水準の F0 値が持続される変種が観察されるが、今回の分析ではこの変種と通常の変種との区別はしなかった。その結果、H% (21794 件)、HL% (6820 件) の計 28614 件の BPM が分析対象となった。

3.2 先行アクセント核数の計算法

先行アクセント核数は、当該 BPM に先行するアクセント核で、かつ当該 BPM が生じた IP に含まれるアクセント核の数と定義した。たとえば、[(ナヲヤワ)] [(ダト)] (ノミヤデ) (ノンダノ) 「直也は誰と飲み屋で飲んだの?」 (“[]” は IP 境界を“()” は AP 境界を表す) という発話の末尾に BPM が生じた場合、その BPM の先行核数は 3 となる。

3.3 話し方のスタイル

概して CSJ の「学会講演」は比較的あらたまったスタイルに特徴付けられ、「模擬講演」は比較的くだけたスタイルに特徴付けられる (前川 2011)。本研究では「学会講演」と「模擬講演」の別、すなわちレジスターの別をスタイルを表す変数として代用する。

3.4 韻律句における BPM の位置

BPM を有する AP が IP の終末に位置するか否かを、韻律句における BPM の位置を表す変数 (IP 非終末と IP 終末の 2 水準) として用いた。BPM を有する AP が IP の終末に位置するか否かは、X-JToBI の Break Indices (BI) のラベルを参照することで一義的に決定できる。

3.5 F0 値の抽出

F0 値は CSJ の XML ファイルに記録されている F0 値 (Hz) を利用した。原則として X-JToBI のトーンラベルひとつにひとつの F0 値が与えられているが、母音の無声化等の理由により F0 値が与えられていない場合もある。このような欠損値は分析対象から除外した。性差や個人差の影響を最小限にするために、F0 値は談話ごとに Z スコアに変換した。

BPM のピッチレンジに関する F0 値の抽出方法は BPM の種類ごとに異なる。H% の場合は上昇の起点 (「上昇起点」と呼ぶ) と上昇の終点 (「頂点」と呼ぶ) の値を計測した。上昇起点として [L%] の持つ値を採用し、頂点として [H%] の持つ値を採用した (ラベル付与位置については図 4 参照)。以降、上昇起点と頂点の差を「上昇幅」と言及する。

一方 HL% の場合は、上昇部分の起点 (「上昇起点」と呼ぶ) と終点 (「頂点」と呼ぶ) に加えて下降部分の終点 (「下降終点」と呼ぶ) を計測した。上昇起点として [L%] の持つ値を採用し、頂点として [pH] の持つ値を採用した。下降終点として [HL%] の値を採用する。以降、上昇起点と頂点の差を「上昇幅」、頂点と下降終点の差を「下降幅」と呼ぶこととする。

4. 分析

4.1 先行アクセント核とスタイルの効果

はじめに、BPM のピッチレンジに対する先行アクセント核数の効果がスタイルごとに異なるか否かを検討した。各条件に該当する BPM の件数を表 2 に示す。件数が 100 未満の BPM (表中の網掛け部分) は以後の分析から除外した。H% のピッチレンジの分析結果を図 5 に、HL% のピッチレンジの分析結果を図 6 に要約する。以降、図中のエラーバーはすべて標準偏差を表す。

BPM の種類の違いに関わらず、頂点、上昇起点、下降終点が先行アクセント核数の増加にしたがって単調に低下する傾向がほぼ一貫して観察される。この傾向は学会講演、模擬講演の両レジスターに認めることができる。HL%の下降幅も両レジスターにおいて先行アクセント核数の増加にしたがって反復的に縮小する傾向が観察される。

最も顕著なレジスターの効果は上昇幅に指摘することができる。学会講演では、上昇幅は両 BPM ともに先行核数の増加にしたがって拡大する。この上昇幅の拡大は模擬講演では観察されない。模擬講演における H%の上昇幅は下降傾向にあり、HL%の上昇幅はほぼ一定に保たれる。このレジスター間の差は、先行核数の増加による頂点の低下の程度（ダウンステップの程度）にその原因を求めることができる。模擬講演では頂点も上昇起点も同程度に低下するのに対して、学会講演では上昇起点の低下の程度よりも頂点の低下の程度が緩やかである。

表2 先行アクセント数とレジスターの別から見た BPM の件数.

レジスター	BPM	先行アクセント核数					
		0	1	2	3	4	
学会講演	H%	2203	7102	3613	991	173	23
	HL%	214	846	479	119	38	10
模擬講演	H%	1138	3878	1995	599	102	17
	HL%	634	2544	1457	415	74	12

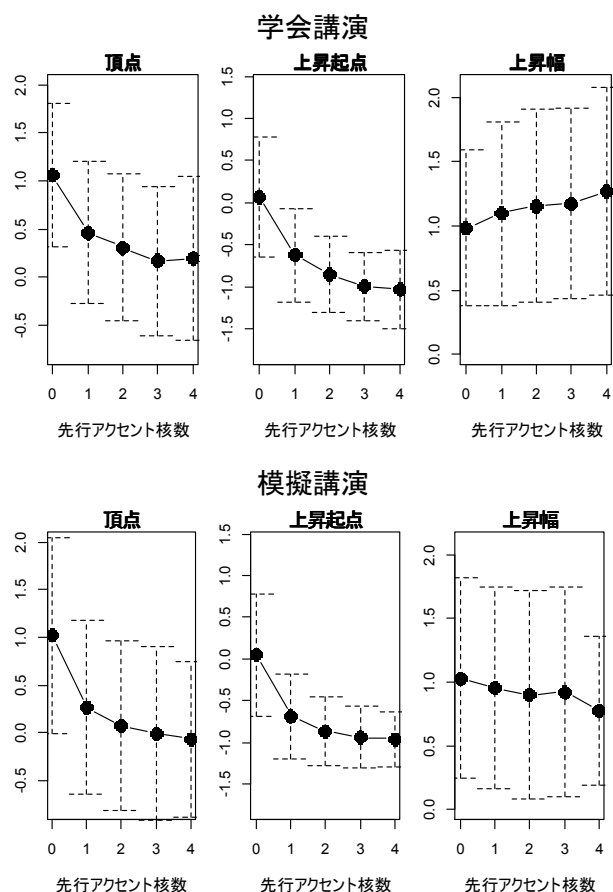


図5 H%のピッチレンジ：レジスター別、学会講演（上）、模擬講演（下）。

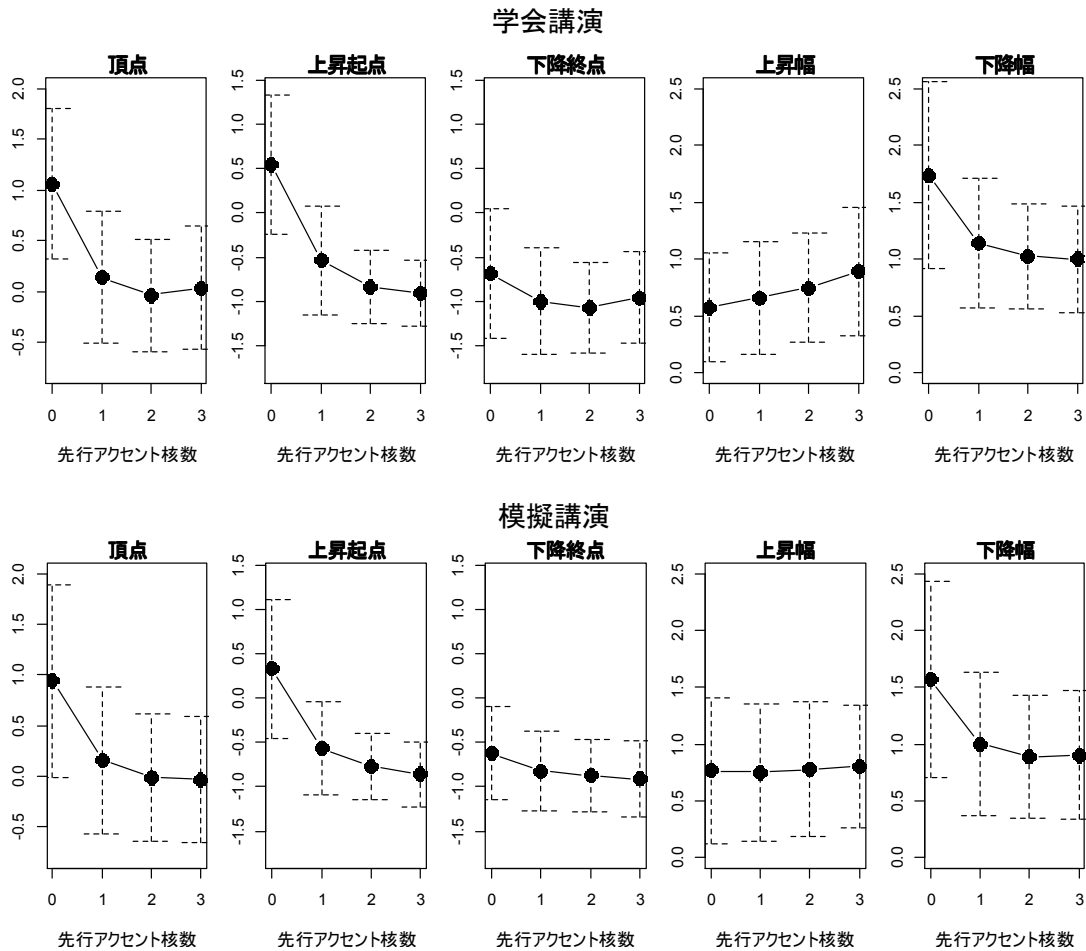


図 6 HL%のピッチレンジ：レジスター別，学会講演（上）、模擬講演（下）。

4.2 IPにおけるBPMの位置の効果

次にIPにおけるBPMの位置（非終末/終末）がBPMのピッチレンジに効果を与えるか否かを検討した。先行アクセント核数とレジスターの双方が影響することがすでに分かっているため、この2つの要因の効果も同時に検討した。各条件に該当するBPMの件数を表2に示す。件数が100未満のBPMは以後の分析から除外した。IPにおけるBPMの位置の別（非終末/終末）だけが異なり、他の要因の値が同一である条件のペアは、両条件におけるBPMの件数が100件以上でない限り、以降の分析から除外した。ただしレジスターが学会講演でかつ先行核数が0である条件のペアは、一方の条件に該当するBPMの件数が100未満であるが、分析対象とした。表中の網掛け部分は分析対象外としたBPMを表す。

表 3 先行アクセント数とレジスターの別から見たBPMの件数。

レジスター	BPM	IP内の位置	先行アクセント核数					
			0	1	2	3	4	5
学会講演	H%	非終末	1098	3061	973	207	37	7
		終末	1105	4041	2640	784	136	16
	HL%	非終末	61	151	32	10	1	0
		終末	153	695	447	109	37	10
模擬講演	H%	非終末	479	1260	335	69	8	2
		終末	659	2618	1660	530	94	15
	HL%	非終末	136	405	77	7	5	0
		終末	498	2139	1380	408	69	12

H%のピッチレンジの分析結果を図7に、HL%のピッチレンジの分析結果を図8に要約する。図から明らかなように、IPにおける位置の効果が観察される。「学会講演」および「模擬講演」における先行核数0のH%の上昇幅、「学会講演」における先行核数0のHL%の下降幅を除いて、すべての特徴（頂点、上昇起点、下降終端、上昇幅、下降幅）は、IP非終末よりIP終末場合の方が低く実現される。この効果はBPMの種類の違い、レジスターの違いに関わらず観察される。

先行核数の効果およびレジスターの効果は4.1で認められたものと同様のものが観察される。すなわち、先行核数が増加するにしたがって、1) 両レジスターにおいてH%とHL%の頂点、上昇起点、下降終端が単調に低下し、2) 両レジスターにおいてHL%の下降幅が縮小するが、3) 模擬講演ではH%、HL%の頂点と上昇起点とが同程度に低下するのに対して、学会講演では上昇起点の低下の程度よりも頂点の低下の程度が緩やかであるため、4) H%とHL%の上昇幅は学会講演においては単調に拡大するが、模擬講演においては拡大することはない。これらの効果は、IPにおけるBPMの位置の違いに関わらず観察される。

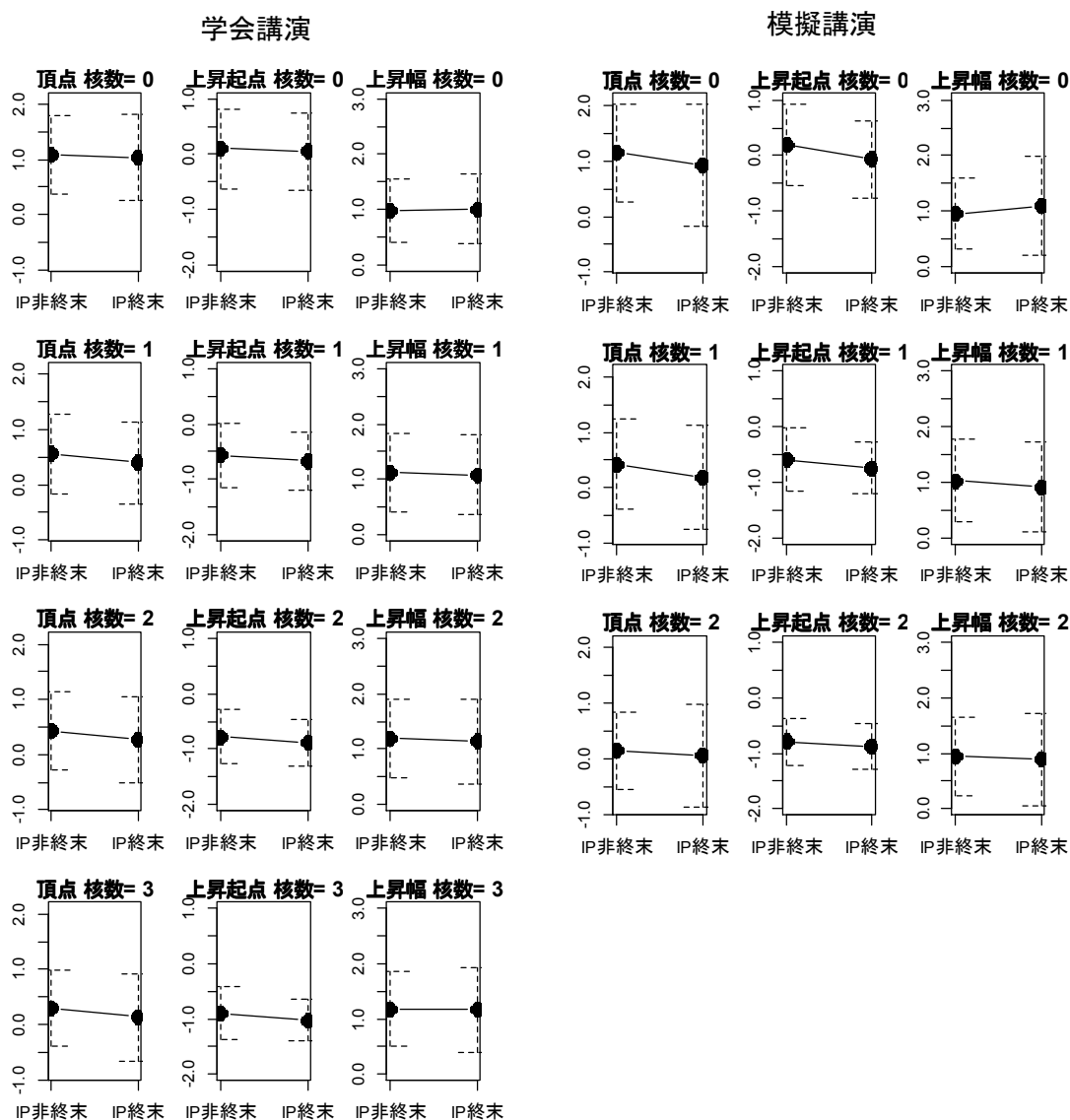


図7 H%のピッチレンジ: スタイル別およびBPMのIP内の位置(IP非終末/IP終末)別。学会講演(左)、模擬講演(右)。「核数」は先行アクセント核数を意味する。

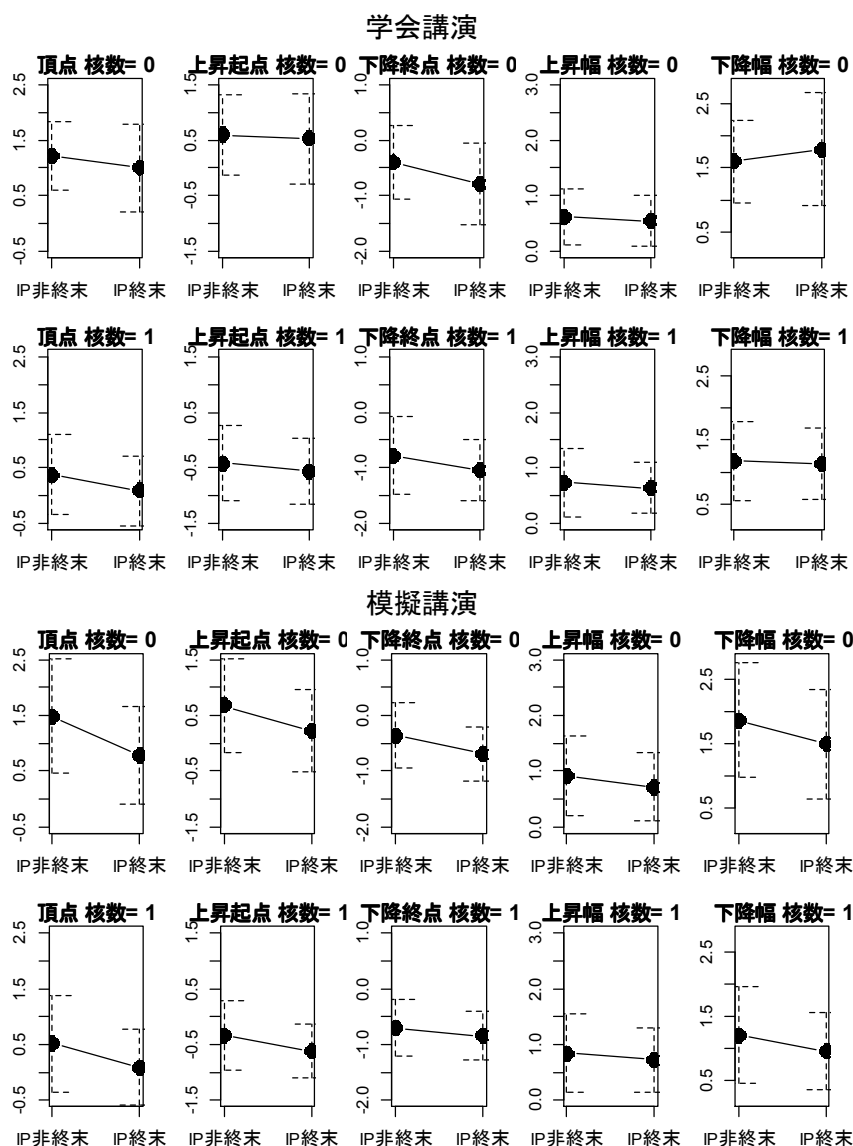


図8 HL%のピッチレンジ：スタイル別およびBPMのIP内の位置（非終末/終末）別。学会講演（上）、模擬講演（下）。「核数」は先行アクセント核数を意味する。

5. 考察と結論

今回の分析結果によって、話し方のスタイルと、韻律句における BPM の位置の双方が、BPM のピッチレンジに効果を与えることが明らかになった。スタイルの差は、ダウンステップの抑制として現れる。「学会講演」「模擬講演」の双方において、先行核数が増えるにつれて、BPM の上昇起点および頂点が低下するが、模擬講演では上昇基点も頂点も同程度に低下するのに対して、学会講演では上昇起点の低下の程度よりも頂点の低下の程度が緩やかである。その結果、「学会講演」では先行核数の増加とともに上昇幅が単調に拡大する。このダウンステップの抑制は、IP における BPM の位置の違いに関わらず観察される。

前川（2011）は、「学会講演」の発話者の一部は、講演前に周到的な練習を重ね発話すべき内容を大部分暗記しているため、通常よりも顕著に長い発話のプランニングを行っている可能性があり、このことが特定のイントネーションパターン（具体的には IP 非終末に現れる BPM）が「学会講演」に頻出する原因となっている可能性を指摘している。BPM のダウン

ステップが抑制される理由も発話のプランニングに見つけられるかもしれない。

BPM のダウンステップが抑制されうるという事実は、日本語のピッチレンジ制御を扱う適切なモデルを構築するための示唆を与える。我々がこれまで明らかにしてきたように、BPM の頂点は、それに先行するアクセント核によって反復的に低下するが、この事実は、AP 主要部のピッチレンジ制御と BPM のピッチレンジ制御が密接に結びついていることを示している。その一方で、先行アクセント核による BPM の頂点の低下が、あるスタイルにおいて抑制されうるという事実は、第 1 に BPM のピッチレンジ制御が AP 主要部のピッチレンジ制御からある程度独立していることを意味し、第 2 に BPM のピッチレンジはスタイル等の要因によって比較的自由に（そしておそらくは連続的に）変動させられうるとを意味する。今後は、BPM をも組み込んだピッチレンジの理論を構築することが必要である。

韻律句における BPM の位置の効果は、（先行アクセント核数が同じにも関わらず）IP 末の AP に生じる BPM の F0 が、IP 末の AP 以前に生じる BPM の F0 より低くなるというものである。この効果はスタイルの違いに関わらず観察される。IP より大きな韻律的単位である発話の終末で F0 が局所的に低下する *final lowering* という現象が知られているが（前川 2011）、IP 末における BPM の F0 低下が *final lowering* と関連した現象なのか否かを検討するのは今後の課題である。また、統語的・意味的な切れ目の強さが BPM の出現率や音声実現に影響することが報告されているので（小磯 2012）、IP 末における BPM の F0 低下を統語的・意味的な切れ目の強さと関連付けて検討することが、BPM のピッチレンジを明らかにするための有望なアプローチのひとつとなると言えるだろう。

付 記

本研究は、国立国語研究所（言語資源研究系）基幹型共同研究「コーパス日本語学の創成」（リーダー：前川喜久雄）および萌芽・発掘型共同研究「会話の韻律機能に関する実証的研究」（リーダー：小磯花絵）による補助、および文部科学省科学研究費補助金・若手研究（B）「イントネーションの音韻論と音声学を峻別する実験手法の確立」（研究代表者：五十嵐陽介）による補助を得ています。

文 献

- 五十嵐陽介、菊池英明、前川喜久雄（2006）「韻律情報」『国立国語研究所報告集 124：日本語話し言葉コーパスの構築法』, pp. 347-453, 国立国語研究所
- 五十嵐陽介、小磯花絵（2012）「日本語話し言葉コーパスにおける句末境界音調のピッチレンジ制御」『第 1 回コーパス日本語学ワークショップ予稿集』, pp. 355-364.
- 川上 稔（1963）「文末などの上昇調について」『国語研究』16, pp. 25-46.
- 小磯花絵（2012）「日本語話し言葉コーパスを用いた複合境界音調の発言継続表示機能の検討」『第 2 回コーパス日本語学ワークショップ予稿集』
- 郡史郎（1997）「日本語のイントネーション型と機能」杉藤美代子（監）、国広哲弥他（編）『日本語音声 2：アクセント・イントネーション・リズムとポーズ』, pp. 169-202, 三省堂
- 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, pp. 111-133.
- 前川喜久雄（2011）「コーパスを利用した自発音声の研究」東京工業大学大学院博士論文
- Igarashi, Yosuke and Hanae Koiso (2012) Pitch range control of Japanese boundary pitch movements, To appear in Proceedings of Interspeech 2012, Portland, Oregon.
- Maekawa, K., H. Kikuchi, Y. Igarashi and J. Venditti (2002) X-JToBI: An extended J_ToBI for spontaneous speech, *Proceedings of the 7th International Conference on Spoken Language Processing*, pp. 1545-1548, Denver, Colorado.
- Pierrehumbert J. and M. Beckman (1988) *Japanese Tone Structure*, Cambridge: MIT press.
- Venditti, J. (2005) The J_ToBI model of Japanese intonation, In S.-A. Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*.172-200, New York: Oxford Univ. Press.
- Venditti, J., K. Maekawa, and M.E. Beckman (2008). 'Prominence marking in the Japanese intonation system', in S. Miyagawa, and M. Saito (eds.), *Handbook of Japanese Linguistics*, pp. 456-512, New York: Oxford University Press.