

自発発話におけるイントネーション句単位の F0 変動の特徴

石本 祐一 (国立情報学研究所 情報学プリンシプル研究系/音声メディアグループ) †

小磯 花絵 (国立国語研究所 理論・構造研究系)

F0 Characteristics at the Level of Intonational Phrase in Spontaneous Japanese

Yuichi Ishimoto (Principles of Informatics Research Division/Speech Media Group, NII)

Hanae Koiso (Dept. Linguistic Theory and Structure, NINJAL)

1. はじめに

小磯・石本 (2012) および石本・小磯 (2012) では、自発音声の発話の韻律的特徴を探るため、イントネーション句 (IP) を単位として発話の基本周波数 (F0) の変動を調べ、発話 (文) 内に強い節境界が存在する場合と存在しない場合において、

- 発話中に強い統語的節境界が存在しない場合 (一つの節で発話が構成される場合)
 1. IP 全体の F0 最大値・最小値が発話内で徐々に下降する
 2. 発話の長さ (IP 数) に関わらず、IP はほぼ一定の高さで始まり一定の高さで終わる (つまり発話の長さによって F0 下降の傾きが異なる)
 3. ただし発話が長い場合、若干高い F0 で発話を開始する
 4. 発話末に著しい F0 下降がみられる (final lowering)
- 発話中に強い統語的節境界が存在する場合 (二つ以上の節で発話が構成される場合)
 1. 発話中の節境界で IP の F0 下降傾向が途切れてリセットされる
 2. F0 下降のリセット時、IP の F0 最小値は発話末のレベルにまで達せず、final lowering に相当する著しい F0 下降はみられない
 3. F0 下降のリセット後、IP の F0 最大値は発話冒頭のレベルにまで達することもあればそれより低いこともある

という傾向がみられることを指摘した。

Pierrehumbert & Beckman (1988) は、アクセント句 (AP) や IP より上位に「発話 (Utterance)」という単位を設定し、その範囲で F0 declination (発話に要する時間の関数として単純に F0 が低下する現象) が、その末尾で final lowering (平叙文末尾で F0 が局所的に下降し発話の終了を示す現象) が生じるとしている。上述の F0 下降傾向は declination のようにも見えるが、発話の長さによって傾きが異なることから、単純に declination とみなすことはできない。また発話末を特徴付ける final lowering が観察されない発話中の強い統語的な節境界において、F0 下降傾向がリセットされることから、「発話」に見られる F0 declination とは解釈しづらい。

† ishimoto@nii.ac.jp

一方 Kawahara and Shinya (2008) は、発話とダウンステップの生じる領域（上記 IP に相当）との間に統語的な節に相当する別の韻律階層が日本語にも存在することを、読み上げ音声を対象とした分析に基づき主張している。F0 下降傾向が節の範囲で観察されるという我々の結果はこの説を支持するものとも解釈できる。しかし、強い節境界で必ず F0 下降のリセットが生じているわけではないこと、強い節境界以外の場所でも F0 下降のリセットが生じる可能性があることなどを考えると、F0 下降を指定する領域として、発話と IP の間に節に相当する単位を設けると結論付けることは拙速であり、さまざまな要因も検討する必要がある。

その一つとして、息などの生理的な要因が挙げられよう。F0 の下降は、一息で発話される範囲において残りの息の量との関係で生じる可能性があるが、意味的なまとまりを示す節境界と息継ぎの位置がある程度一致していたために、結果として統語的な節境界で F0 下降のリセットが観察されたとも考えられる。そこで本研究では、節境界だけではなく、息継ぎ（によって生じるポーズ）や発話時間などを含む物理的な特徴量を総合的に考慮した分析を行い、発話中の IP 単位の F0 変動に関わる要因について多角的に検討する。

2. データ

2.1 コーパス

分析には『日本語話し言葉コーパス (*Corpus of Spontaneous Japanese*:以下 CSJ)』（前川 2004）を用いた。CSJ は自発性の高いモノログを中心に構成された話し言葉コーパスであり、学会における口頭発表（以下「学会講演」と、一般話者による主に個人的な内容に関するスピーチ（以下「模擬講演」）を主対象としている。実際の分析には CSJ 第 3 刷に基づき作成された RDB（小磯ほか 2012）を用い、CSJ 全体にあたる 661 時間の音声のうち、「コア」と呼ばれるデータ範囲の中から学会講演 70（約 19 時間）・模擬講演 107（約 20 時間）を分析対象とした。

2.2 節単位

CSJ に付与されている節単位情報（丸山ほか 2006）を分析に利用する。節単位は原則「節（clause）」の境界によって得られる文法的・意味的なまとまりを持った単位であり、節境界の構造的な切れ目の大きさの観点から以下の 3 つに分類される。

絶対境界（Absolute boundary） いわゆる文末に相当する境界。

強境界（Strong boundary） 後続の節に対する従属度の低い、切れ目の度合いが強い節境界。

弱境界（Weak boundary） 後続の節に対する従属度の高い、切れ目の度合いが弱い節境界。

これらの境界は形態素解析結果に基づき自動で判別され、人手による修正・操作を経た上で、絶対境界、強境界のいずれかで区切られる単位が節単位と認定されている。

2.3 発話・イントネーション句・アクセント句

Pierrehumbert & Beckman (1988) の韻律理論では、「発話ーイントネーション句 (IP) ーアクセント句 (AP)」という韻律的階層構造が仮定される。AP は、第 1 モーラから第 2 モーラ付近にかけての F0 の上昇と句末への緩やかな下降を有し、かつアクセント核による下降を最大ひとつ持ちうる単位と定義される。IP は AP の上位階層に位置し、AP のピッチレンジを指定す

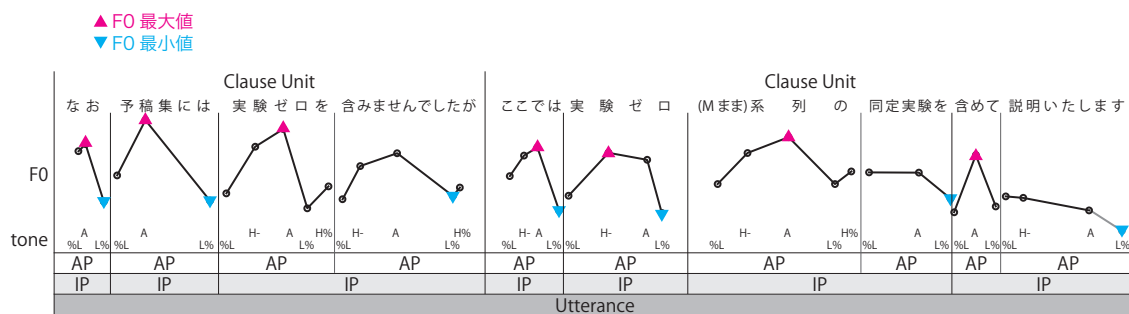


図1 韻律的階層構造

る単位と定義される。アクセント核が引き起こす後続 AP のピッチレンジの縮小効果は、IP の範囲で観察される。発話は上述の通り、IP の上位階層に位置し、F0 declination が観察され、かつその末尾で final lowering が生じる範囲と定義される。以上をまとめると、各階層で F0 は、

- AP の領域では、山状に上昇後下降する
- IP の領域では、AP 単位で低下する（ダウンステップ）
- 発話の領域では、発話末に向かって低下し、発話末で急激に低下する

といった変動を示す。この韻律的階層構造を図1に示す。

本研究では IP を単位とした発話の F0 変動の特徴を探る。CSJ にはラベリングスキーム X-JToBI (五十嵐ほか 2006) に基づき韻律情報が付与されているが、この中に、韻律境界の切れ目の強さに関する情報として Break Index (BI) が存在する。BI=2 は AP 境界、BI=3 は IP 境界、BI=F はフィラー境界、BI=D は言い淀み境界に対応する。ここでは、BI=3 で区切られる範囲を IP と認定し、フィラー、言い淀み部分を除いて分析に用いた。ただし、フィラーを狭んでダウンステップが続く場合はフィラーを内包する形で IP を認定した。

このように IP を認定した上で、IP の F0 特徴量として X-JToBI に基づく Tone 情報から、
 F0 最大値 IP 頭の AP の句頭音調 (H-) あるいはアクセント核 (A) のうち高い方の F0 値
 F0 最小値 IP 末尾の AP の下降音調 (L%) の F0 値
 を求めた。また、分析において性差・個人差の影響を小さくするために、F0 値は話者ごとの平均 F0・標準偏差によって Z スコアへ変換して用いる。

韻律的な発話の境界情報は X-JToBI では示されないため、本研究では発話に相当する単位を、前節に示した絶対境界によって区切られる区間とする*1。分析対象となる発話数は 9,885 となった。

3. IP の F0 最大値・最小値と IP 前後のポーズの関係

3.1 方法

本節ではまず、IP 前後のポーズと F0 最大値・最小値との関係について調べる。

1 節でもまとめたように、小磯・石本 (2012) および石本・小磯 (2012) では、

1. IP の F0 最大値・最小値が発話末へ向けて徐々に低下する

*1 明示的な文末表現が置かれるもののほか、「と文末」や「体言止め」なども含む。

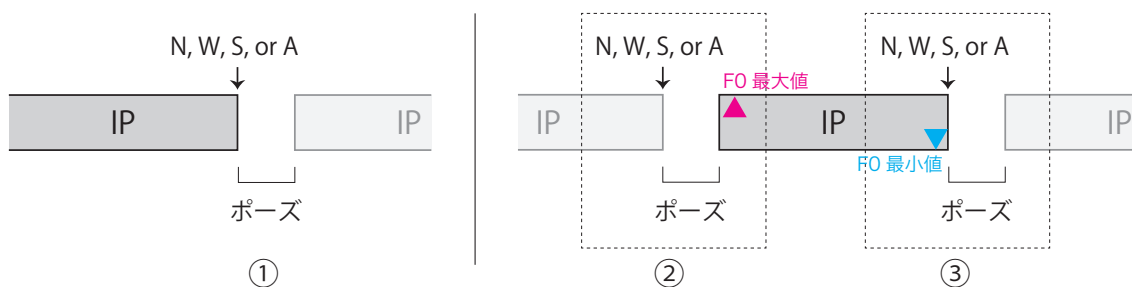


図2 IP前後のポーズと節境界（非境界:N, 弱境界:W, 強境界:S, 絶対境界:A）

表1 IP前後の節境界とIP数（非境界:N, 弱境界:W, 強境界:S, 絶対境界:A）

IP 前	IP 後				合計
	N	W	S	A	
N	23706	4326	1841	920	30793
W	2793	545	239	141	3718
S	1263	119	40	14	1436
A	1227	135	50	26	1438
合計	28989	5125	2170	1101	37385

2. 発話中に強い節境界が見られる場合、IPのF0下降はそこでリセットされるという傾向が観察された。このように、先の分析結果からは統語的な境界である強境界において下降のリセットという急激なF0変動が起こっているように見えるが、実際は統語的な境界が主な要因ではなく、強境界（および絶対境界）で生じやすい息継ぎやポーズ（空白期間）がF0下降の途切れの要因となっている可能性が考えられる。

そこでIPの前後のポーズによりF0最大値・最小値に違いが現れるか分析を行う。具体的には、IP末の節境界とその直後のポーズの関係（図2①）を確認した上で、IP前のポーズとF0最大値（図2②）およびIP後のポーズとF0最小値（図2③）の関係を調べる。節境界の種類は、2.2節に示す「絶対境界」「強境界」「弱境界」およびこれら節境界の存在しない「非境界」の4種とする。なお、CSJには息継ぎのタイミングは記録されていないため、ポーズをその代替的指標として取り上げる。

ポーズの長さが2秒以下のIPに限定したところ、分析対象となるIPの数は37,385であった。IP前後の節境界の種類とIP数を表1に示す。

3.2 結果と考察

IP末の節境界の種類とその直後のポーズ長との関係について、図3に示す。分散分析の結果、節境界の違いは有意であった（ $F(3.000, 3394.846) = 1123.324, p < 0.001$ ）。Tukey法による多重比較によれば、すべての節境界間に有意差があった。予想していたように、統語的な切れ目が強いほどポーズが長くなるという傾向が顕著に見られる。

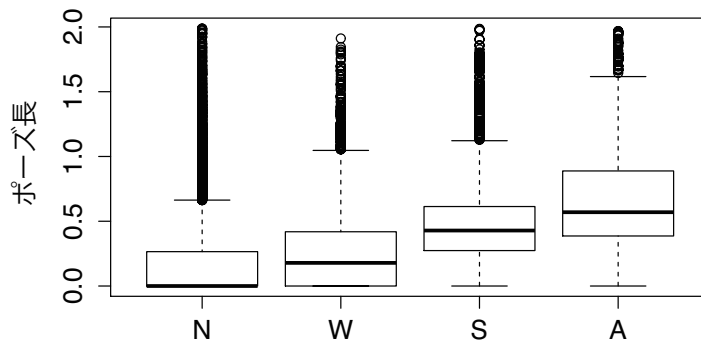


図3 IP 末の節境界の種類とその直後のポーズ長 (非境界:N, 弱境界:W, 強境界:S, 絶対境界:A)

次に IP 前のポーズの長さ と IP の F0 最大値 (図2②) の散布図を、IP 直前の節境界の種類ごとに図4に示す。ポーズが F0 下降のリセットと関わるならば、節境界の種類に関わらず、ある程度の長さのポーズが置かれると、そこで F0 下降はリセットされ、直後の IP の最大値は高くなることが予想される。しかし図4からわかるように、非境界と強境界ではその傾向は全く見られず、また弱境界と絶対境界についてもわずかにその傾向がうかがえる程度である。

IP 後のポーズの長さ と IP の F0 最小値 (図2③) の散布図を図5に示す。IP 単位の F0 下降がポーズの位置で止まるのであれば、ポーズの前の F0 最小値はポーズがない場合よりも低くなるはずである。図5からわかるように、確かに統語境界の種類によらずこの傾向は観察されるものの、その効果は決して大きくはない。

このようにポーズの長さだけに着目すると、IP の F0 値に与える影響は部分的かつ微小であることが分かった。

4. IP の F0 最大値・最小値と発話の時間パラメータの関係

4.1 方法

本節では、ポーズ以外の発話の物理量も取り上げ、F0 最大値・最小値との関係を調べる。具体的には、発話の時間に関わるパラメータとして、

- IP 直前のポーズ長
- IP 直後のポーズ長
- 当該 IP を含む発話の長さ
- 当該 IP を含む節単位の長さ
- IP の長さ
- IP の開始時刻から発話末までの時間
- IP の開始時刻から節単位末までの時間

を話者ごとに Z スコアに変換して説明変数とし、F0 最大値および F0 最小値を応答変数とする一般化線形混合モデルを構築した。なお、話者の違いによる影響を誤差項としてモデルに加えてある。分析対象は前節の分析と同じ IP である。F0 最大値については IP 直前の節境界の種類で、F0 最小値は IP 直後の節境界の種類でグループ分けを行った上でモデルに当てはめた。

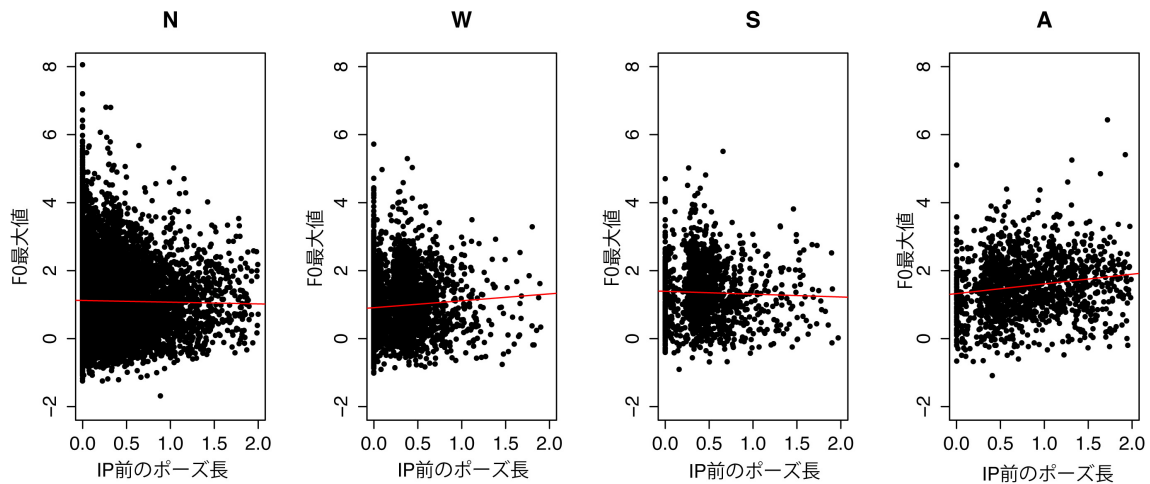


図4 IP直前のポーズの長ささとF0最大値（図上部のN, W, S, AはIP直前の節境界）

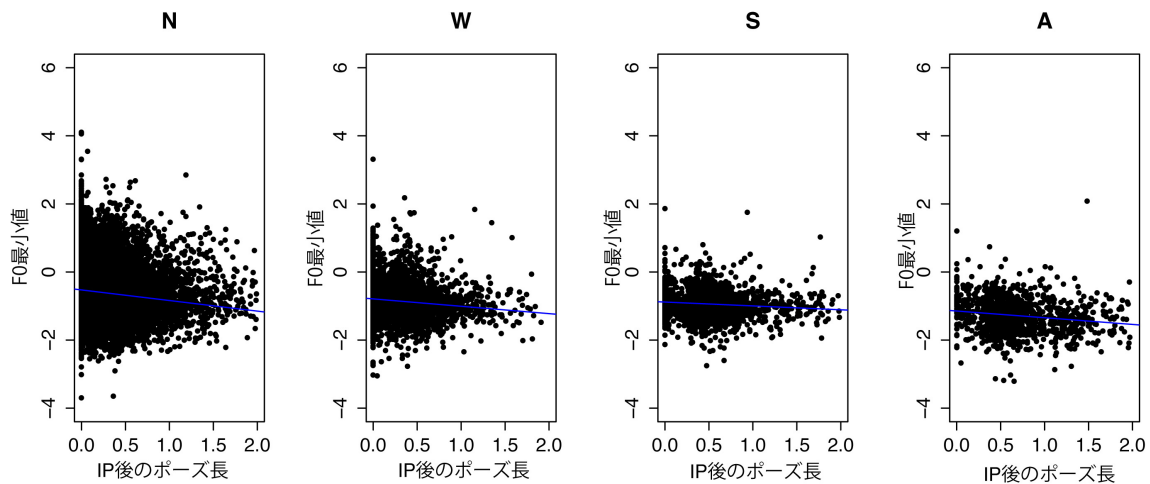


図5 IP直後のポーズの長ささとF0最小値（図上部のN, W, S, AはIP直後の節境界）

4.2 結果と考察

4.2.1 IP直前の節境界とF0最大値

IP直前の節境界ごとのF0最大値に対する結果を表2に示す*2。なお、直前が強境界および絶対境界の場合の節単位末までの時間は節単位長に等しく、加えて絶対境界の場合は発話末までの時間と発話長も等しくなるため、それぞれ除外している。

表2から、すべての節境界に共通して、IP長が長いほど当該IPのF0最大値は高くなる傾向にあることが分かる。この結果は、IP内のAPのダウンステップでF0が下がりにすぎないよう、長いIPの場合はIP冒頭を高めF0で始めるといった調整が、場所を問わず共通して見られることを意味する。

*2 一般化線形混合モデルの構築にはRのlme4パッケージに収録されているlmer関数を用いた。また、p値はlanguageRパッケージのpvals.fnc関数で算出した。

表2 一般化線形混合モデルの説明変数（応答変数: F0 最大値）

	IP 直前が非境界 (N)				IP 直前が弱境界 (W)			
	Estimate	Std. Error	t 値		Estimate	Std. Error	t 値	
(Intercept)	1.012	0.018	56.99	p<0.01 **	0.909	0.022	40.86	p<0.01 **
前ポーズ長	0.012	0.006	2.03	p=0.04 *	0.053	0.015	3.43	p<0.01 **
後ポーズ長	-0.074	0.006	-12.65	p<0.01 **	-0.115	0.016	-7.03	p<0.01 **
発話長	-0.045	0.008	-5.84	p<0.01 **	-0.010	0.020	-0.50	p=0.61
節単位長	-0.197	0.008	-24.76	p<0.01 **	-0.053	0.022	-2.43	p=0.02 *
IP 長	0.282	0.006	51.33	p<0.01 **	0.285	0.016	17.53	p<0.01 **
発話末まで	0.048	0.008	6.11	p<0.01 **	-0.004	0.021	-0.19	p=0.85
節単位末まで	0.221	0.008	28.06	p<0.01 **	0.124	0.022	5.55	p<0.01 **
	IP 直前が強境界 (S)				IP 直前が絶対境界 (A)			
	Estimate	Std. Error	t 値		Estimate	Std. Error	t 値	
(Intercept)	1.316	0.038	34.69	p<0.01 **	1.32	0.051	26.07	p<0.01 **
前ポーズ長	0.019	0.020	0.92	p=0.36	0.104	0.014	7.154	p<0.01 **
後ポーズ長	-0.043	0.028	-1.56	p=0.12	-0.040	0.025	-1.58	p=0.11
発話長	-0.032	0.032	-1.00	p=0.32	0.064	0.029	2.18	p=0.03 *
節単位長	0.044	0.030	1.44	p=0.15	-0.001	0.028	-0.04	p=0.97
IP 長	0.317	0.023	13.95	p<0.01 **	0.283	0.021	13.55	p<0.01 **
発話末まで	0.037	0.035	1.06	p=0.29				
節単位末まで								

** : 1% 水準, * : 5% 水準, + : 10% 水準

IP 直前が非境界の場合、上記に加えて次の傾向が見られる。まず、節単位末までの時間が長いほど、つまり節の前部に位置するほど、F0 最大値は高い傾向にある。この結果は、非境界の多くで単純に F0 が下降する傾向が見られることを意味する。発話末までの時間も同じく有意であることから、発話の残りの時間が長いほど F0 が高いという、発話全体を通した declination の存在もうかがえる。節単位長や発話長の影響もみられるが、節単位や発話が長ければそれだけ IP 単位の下降の頻度が増すため、F0 が下がり切らないよう下がり幅（下降の傾き）を調整していることを表していると考えられる。また、IP 直前のポーズが長いと F0 最大値が高くなるという傾向も見られる。この結果は、非境界の多くは下降傾向を見せるものの、長いポーズが見られる場合には、そこで IP 単位の F0 下降がリセットされる可能性があることを示唆する。

IP 直前が弱境界の場合も、非境界とほぼ同様の傾向が見られるが、非境界で有意であった発話長・発話末までの時間の効果は見られない。非境界も弱境界も、長いポーズのあとでは F0 下降はリセットされる傾向が示唆されるが、図3に示したように、そもそも非境界ではポーズが置かれることは少なく（中央値 0）、結果として F0 下降がリセットされることも相対的に少ないと考えられる。その結果、統計的な観点から、発話全体を通した declination の存在が非境界にのみ見られたものと推測される。

IP 直前が強境界の場合（当該 IP が節単位冒頭に位置する場合）、非境界や弱境界とは異なり、その上位階層に位置する単位（強境界の場合は発話）の長さや単位末までの時間の影響は

表3 一般化線形混合モデルの説明変数 (応答変数: F0 最小値)

	IP 直後が非境界 (N)					IP 直後が弱境界 (W)				
	Estimate	Std. Error	t 値	p	Signif.	Estimate	Std. Error	t 値	p	Signif.
(Intercept)	-0.582	0.015	-39.72	p<0.01	**	-0.787	0.017	-45.63	p<0.01	**
前ポーズ長	0.023	0.004	6.22	p<0.01	**	0.017	0.007	2.31	p=0.02	*
後ポーズ長	-0.089	0.005	-17.10	p<0.01	**	-0.063	0.008	-7.74	p<0.01	**
発話長	-0.022	0.006	-3.74	p<0.01	**	-0.010	0.010	-0.94	p=0.35	
節単位長	-0.092	0.006	-15.07	p<0.01	**	-0.029	0.010	-2.80	p<0.01	**
IP 長	-0.159	0.004	-35.69	p<0.01	**	-0.060	0.007	-8.19	p<0.01	**
発話末まで	0.035	0.006	5.72	p<0.01	**	0.024	0.010	2.29	p=0.02	*
節単位末まで	0.111	0.006	17.92	p<0.01	**	0.039	0.011	3.65	p<0.01	**
	IP 直後が強境界 (S)					IP 直後が絶対境界 (A)				
	Estimate	Std. Error	t 値	p	Signif.	Estimate	Std. Error	t 値	p	Signif.
(Intercept)	-0.873	0.025	-35.57	p<0.01	**	-1.177	0.039	-30.49	p<0.01	**
前ポーズ長	0.004	0.009	0.43	p=0.67		0.0007	0.015	0.05	p=0.96	
後ポーズ長	-0.042	0.008	-5.40	p<0.01	**	-0.045	0.011	-4.13	p<0.01	**
発話長	-0.010	0.012	-0.82	p=0.41		-0.020	0.019	-1.06	p=0.29	
節単位長	-0.019	0.011	-1.68	p=0.09	+	0.0004	0.018	0.02	p=0.98	
IP 長	-0.034	0.008	-4.34	p<0.01	**	0.006	0.013	0.50	p=0.61	
発話末まで	0.013	0.011	1.20	p=0.23						
節単位末まで										

** : 1% 水準, * : 5% 水準, + : 10% 水準

見られないことから、強境界をはさんで F0 は単純に下降せずリセットする傾向にあること、またそのリセット後の高さが発話の長さや発話中の位置に因らないことを意味する。また強境界では IP 直前のポーズ長の影響は見られない。つまり、強境界ではポーズの長さに関わらず F0 下降のリセットが生じる可能性が高いことを意味する。

IP 直前が絶対境界の場合 (当該 IP が発話冒頭に位置する場合)、発話長が長いほど F0 最大値が高い傾向にあることから、長い発話単位の場合、末尾で F0 が下がり切らないよう発話の冒頭をより高い F0 で開始する傾向にあることが示唆される。また絶対境界の場合、強境界とは異なり IP 直前のポーズ長の影響が見られる。発話末では通常ポーズが置かれるが、まれに次の発話冒頭 (接続詞など) までポーズを置かず連続して発話することがある。そのような場合、F0 下降は発話冒頭要素まで続き、そのあとにリセットが生じる可能性もある。あるいは、上述の通りより長い発話をする場合はより高い F0 で発話を開始するが、長い発話のプランニングのためにポーズが相対的に長くなった結果、ポーズ長と発話冒頭の F0 最大値との間に相関が見られた可能性もある。

4.2.2 IP 直後の節境界と F0 最小値

IP 直後の節境界ごとの F0 最小値に対する結果を表 3 に示す。IP 直後の節境界が絶対境界/強境界の場合とは、当該 IP は発話/節単位の末尾に位置することを意味する。なお、直後が強境界および絶対境界の場合の節単位末までの時間は IP 長に等しく、加えて絶対境界の場合

は発話末までの時間も IP 長に等しくなるため、それぞれ除外している。

結果から、絶対境界以外では IP 長の影響がみられ、IP が長いほど末尾の F0 は低い値となることがわかる。前節の分析で、IP 内の AP のダウンステップで F0 が下がりすぎないように、長い IP では冒頭を高めめの F0 で始めるといった調整がとられる可能性を指摘したが、その調整だけで IP 末尾の F0 値が一定に納まるわけではなく、長い IP では F0 がより低い位置まで低下することを意味する。絶対境界直前（発話末）で IP 長の影響が見られないのは、発話の F0 の下限が固定的であるという小磯・石本（2012）の結果と整合的である。当該話者の発話のピッチレンジの下限よりも低くなることはなく、下げ止まりの状態になるということであろう。

IP 直後が非境界と弱境界の場合、IP 長と後続ポーズ以外の結果は、前節の F0 最大値と同じ傾向を示しており、同様の考察を導くことができる。つまり、非境界および弱境界ではポーズによって F0 下降のリセットが生じることはあるものの、総じて F0 下降が続く位置と言える。

IP 直後が強境界の場合（当該 IP が節単位末尾に位置する場合）、前節の F0 最大値と同様に発話末までの時間の影響がないことから、強境界をはさんで F0 は単純に下降せずリセットする傾向にあることが示唆される。また節単位長が長くなると F0 最小値はより低くなるという傾向も見られる。強境界の場合、リセット直前の F0 最小値は発話末ほどは下がらず、また値も一定しないが（小磯・石本 2012）、リセット直前の F0 最小値には節単位の長さに関わる可能性が示唆される。

IP 直後が絶対境界の場合（当該 IP が発話末尾に位置する場合）、ポーズ長以外の時間のパラメータの影響を全く受けていないことから、前述のように発話末の F0 の下限はかなり固定的であることが分かる。

最後に、すべての節境界に共通して IP 直後のポーズ長が有意であり、直後に長いポーズがあると F0 最小値は低くなる傾向にあることがわかる。強境界の場合、上述の通り節単位長が長いほどリセット直前の F0 最小値はより低くなる傾向にある。仮に長い節単位ほど末尾での息の残量が少なくなり、直後により長い息継ぎの時間（ポーズ）を置く傾向があるとするならば、節単位の長さを媒介に IP 直後のポーズ長と直前の F0 最小値との間に相関が見られたと考えることができる。その他の境界の場合も、同様の可能性が考えられる。

5. おわりに

本稿では、小磯・石本（2012）および石本・小磯（2012）で観察された IP 単位の F0 下降現象と強境界での下降のリセットが統語的な要因だけに依存しているのかを調べるため、息継ぎ（ポーズ長）や発話時間などの物理的な特徴量を含めた総合的な分析を行った。その結果、小磯・石本（2012）および石本・小磯（2012）で観察された各種傾向が改めて統計的に確認されたほか、新たに次のことが明らかになった。

1. 非境界・弱境界では通常、発話（節単位）末に向けて F0 が低下する傾向が見られるが、長いポーズが置かれると、ポーズ前で F0 が低下しポーズ後に高い F0 で始まる「F0 下降のリセット」が生じる可能性がある。
2. 強境界では、明らかに単純な F0 下降とは異なる F0 変動をしており、F0 下降のリセットが起きていると考えられるが、F0 最大値にはポーズの影響がみられないことから、

ポーズの長さによらず F0 下降のリセットが生じている可能性が高い。強境界のように発話内容の意味的なまとまりの強い区切りでは、ポーズの有無といった生理的要因に関わらずリセットが生じることが示唆される。

3. 節境界で F0 下降のリセットが生じる場合、リセット直前の F0 最小値は発話末ほどは下がらずまた値も一定しないが、この F0 最小値に節単位長が関わる可能性がある。
4. 発話が長いほど発話冒頭の F0 最大値が高くなるのと同様に、IP が長いほど IP 冒頭の F0 最大値が高くなることから、発話・IP という大小異なる単位で、発話前にプランニングをして各单位中で F0 が必要以上に下がりきらないような調整が行われている可能性がある。

このように、発話に見られる IP 単位の F0 下降やそのリセットに関して多くの事実が徐々に明らかになってきたが、発話中に見られる F0 下降のリセットを引き起こす要因や発話全体に見られる declination との関係などについてはまだ解明されていない。この点については今後の課題としたい。

参 考 文 献

- 小磯花絵, 石本祐一 (2012) 「日本語話し言葉コーパスを用いた「発話」の韻律的特徴の分析 – イントネーション句を切り口として –」第 1 回コーパス日本語学ワークショップ予稿集, pp. 167–176.
- 石本祐一, 小磯花絵 (2012) 「日本語話し言葉コーパスを用いた統語境界におけるイントネーション句変動の分析」第 2 回コーパス日本語学ワークショップ予稿集, pp. 239–246.
- Pierrehumbert, Janet B. and Mary E. Beckman (1988) *Japanese tone structure*, Cambridge: MIT Press.
- Kawahara, Shigeto and Takahiro Shinya (2008) “The intonational of gapping and coordination in Japanese: evidence for intonational phrase and utterance,” *Phonetica*, 65, pp. 62–105.
- 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.
- 小磯花絵, 伝康晴, 前川喜久雄 (2012) 「『日本語話し言葉コーパス』RDB の構築」第 1 回コーパス日本語学ワークショップ予稿集, pp. 393–400.
- 丸山岳彦, 高梨克也, 内元清貴 (2006) 「節単位情報」日本語話し言葉コーパスの構築法 (国立国語研究所報告 124), pp. 255–322.
- 五十嵐陽介, 菊池英明, 前川喜久雄 (2006) 「韻律情報」日本語話し言葉コーパスの構築法 (国立国語研究所報告 124), pp. 347–453.

※ 本研究は萌芽・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) による成果である。