

<論文>

## 自発発話音声から見た日本語音調の動態

### Dynamic tone characteristics of Japanese observed in spontaneous utterances

佐藤 大和  
Hirokazu Sato

東京外国語大学大学院総合国際学研究院  
Tokyo University of Foreign Studies

**要旨:** 本論文は、自由に発話された音声資料の音調分析を行い、自発発話における日本語音声の句音調の特性、また規範的な日本語アクセントの型が連続音声の中でどのような動的形態として実現されているか等について分析した結果の報告である。分析資料は、「日本語話し言葉コーパス(CSJ)」から選択された東京方言話者3名による独話資料である。自発発話における句音調には、アクセント句が際立つように発話された卓立型音調様式と単調な発話の非卓立型発話様式があり、必ずしも朗読音声に見られるようなピッチの漸近的下降特性が見られないことを示す。また、アクセント句の音調に関しては、アクセントが知覚される拍(アクセント拍)における“上昇”“下降”“平坦”等の音調形式、これらアクセント音調制御と調音制御の関連、アクセント位置が後続拍にくる「遅下がり」現象の成因等について記述する。これらの分析結果から、アクセント句の音調構造モデルを議論する。

**Abstract:** This paper describes several analyzed results of phrase intonation characteristics and tonal dynamics of Japanese accents in spontaneously uttered Japanese. Speech materials for the analysis are monologue talks by three speakers of Tokyo dialect accent selected from “The Corpus of Spontaneous Japanese”. Two intonation styles observed in spontaneous utterances are firstly shown. One is prominent-type intonation including saliently uttered accentual phrases, and the other is monotonously spoken non-prominent-type intonation. The asymptotically declined intonation, which is typical in read speech of arranged text, is little observed in both intonation styles. Secondly, the paper reports dynamic tone forms of accented syllables such as “rising”, “falling” and “level” tones, the relationship between tonal control and articulatory timing, and causes of the delayed pitch-falling phenomena that the accent position shifts from an accented syllable onto a succeeding syllable. And moreover, a structural model of pitch frequency contour in the accentual phrase of Japanese is discussed.

**キーワード:** 自発発話, 日本語アクセント, 音調動態, 遅下がり, 句音調

**Keywords:** Spontaneous utterance, Japanese accent, Dynamic tone form, Delayed pitch falling, Phrase intonation

#### 1. はじめに

日本語(共通語)のアクセントの型は、単語を構成する拍の「高」「低」の配置、もしくは音調が「高」から「低」へ変化する際の「高」の拍の位置(アクセント核)によって記述される。これらの型は、拍を意識した比較的丁寧な発音もしくはその内省によって把握される。一方、実際の発話における音調の動的特性を観測すると、こうした規範的なアクセントの高低の型がそのまま音調特性上で実現されているわけではなく、アクセントの「遅下がり」現象(杉藤1980)のように、アクセントが認知される拍と音響特性上で観測されるアクセント位置との間に乖離があることも知られている。この研究は、日本語アクセントの型が、実際の連続音声のなかでどのような動態形式として実現されているか、また逆に、特定のアクセント型をもたらす音調の動的特徴は何かを明らかにすることを目的としている。



本稿の著作権は著者が保持し、クリエイティブ・コモンズ 表示 4.0 国際ライセンス(CC-BY)下に提供します。  
<https://creativecommons.org/licenses/by/4.0/deed.ja>

また一方、アクセントやイントネーションなどを担う言語の超分節的特徴あるいは韻律的特徴は、主に音節、語、句等の発話音声の音調特性（ピッチ周波数の時間特性）の観点から研究が行われてきたが、日本語をはじめとしてこれまでの研究は、研究目的に合致するように選択・構成された語や文のリストに基づいて発話された音声資料によってなされてきたように思われる。つまり「テキストの読み上げ音声」や「指示に基づく発話音声」の分析が主体であり、日常の言語生活における自発的かつ自由に発話された音声の音調特性の研究は不十分であった。本論は、日本語の音調特性を自発発話という発話様式上から明らかにすることも一つの目的であるが、特に日本語アクセントに関して、発話の仕方を規定しない自由な発話の方がより多様で顕著な音調動態形式をもたらすのではないかと考え、自発発話音声によってアクセントの音調実現形態を探ることとした。また、複数のアクセントを含み音調的なまとまりのある発話単位についても、自発発話音声特有のイントネーション様式のあることを見出したので併せて報告する。

## 2. 分析資料と分析方法

### (1) 音声資料

分析資料として、国立国語研究所「日本語話し言葉コーパス(CSJ)」の中の東京方言話者3名の独話資料を用いた。これらは、女性2名：話者：FA, FB, 男性1名：話者：MAによる模擬講演音声であり、祖母の思い出、犬の話、趣味の映画などについて自由に語る内容である。各人7～10分程度の時間長の音声である。

これらの音声のピッチ周波数(F0)とともに、各話者の発話段落末の最低ピッチ周波数を基準値とするSemitone(ST)を求め、F0とSTの双方から音調特性と音韻区分、アクセント部分、音調句区分等との関係を分析した。Semitoneは、話者ごとの音域の違いを正規化して比較できる利点があり、本論文ではSTの分析結果を主に示す。Semitoneは以下の式で求められる。

$$\text{Semitone (ST)} = 12 * \log_2 (F0/F0b)$$

ここで、F0bは各話者の発話段落末の基底ピッチ周波数であり、発話全体の音調特性を観測して決められた。F0bの値は、女性話者FAでは135 Hz、話者FBでは140 Hz、男性話者MAでは80 Hzとした。Semitoneは、1オクターブの音域を12段階で表現するスケールとなっており、各発話者ともピッチの変動範囲はおおよそ2オクターブ、すなわち24 ST程度である。

アクセントに関しては、音声データの聞き取りによってアクセント型の判断を行うとともに、ピッチ周波数特性の形に基づいてアクセント位置を定めた。聞き取りによってアクセントがあると定められた拍を「アクセント拍」、音響特性（ピッチ周波数）から定められたアクセントの時間軸上の位置を「アクセント位置」、アクセント句の高さのピーク値を「アクセント句のプロミネンス」と呼ぶことにする。アクセント位置は必ずしもアクセント拍上にはなく、特にアクセント位置がアクセント拍の後続拍上にくる現象は、いわゆるアクセントの「遅下がり」である。

アクセント位置は、CJSのドキュメント（五十嵐・菊池・前川 2006）の記載に準じており、アクセント拍とこれに後続する拍のピッチ周波数パターンに基づいて以下の手法で定められた。

- (a) 上昇ののち下降するピッチ特性の場合はそのピーク時点
- (b) 平坦もしくは緩やかな変化から急峻なピッチ下降がある場合は下降の開始点
- (c) ピッチが下降する特性のみの場合はその開始時点

## (2) 分析の音声区分

分析資料の独話音声は、比較的長いポーズで区切られた発話段落 (Utterance Group) から構成されるが、その段落をまず発話単位 (Utterance Unit: UU) の系列として区分化する。UU は、発話連続を構成するまとまりのある発話区分であって、その音声上の区分設定としては、一定長のポーズ、句末における上昇や上昇・下降の境界音調の存在、フィラーの存在、句末母音の引き延ばし等を参照し、さらに一定の意味的まとまりも考慮して決められた。ひとつの UU には、1 から 4 程度のアクセント句が含まれる。自発発話においては、言いよどみ、発話の未完やそれに続く次への展開などが多く、書き言葉における文法的な句や文の境界のように必ずしも明確に区分化されるものではないが、表現と伝達のため一定の発話のまとまりというものは認められる。こうした発話単位の音調特性を自発発話の「句音調」と呼ぶことにする。話者 FA と MA の場合の、発話段落の発話単位への区分化の例を以下に示す。

話者 FA : ((アノ) はフィラー)

(であたしはもう) UU1, (大きくなるまで) UU2, (というか本当に大人になるまで幸いなことに) UU3, (身近な人との) UU4, ((アノ) 別れというものが無くて) UU5

話者 MA :

(年齢的にも) UU1, (みんな映画をちょこちょこ見だす年齢だと例えば) UU2, (言えるんですけども) UU3

発話単位 UU は、さらに複数のアクセント句から構成される。通常のアクセント句の定義では、たかだか 1 個のアクセントを有する音声単位をいうが、アクセントのある語に付属語や形式名詞などが付随して、これら要素のアクセントも音調上に現れる場合がある。このような副次的アクセントが付いて複数アクセントからなる句もここではアクセント句と呼ぶことにする。例えば、上記の話者 FA の発話単位 UU2 において、「大きくなるまで」(o]okikuna]rumade) のように二つのアクセントで発話される場合があり、このような場合も一つのアクセント句とした。

## 3. 句音調の特性

これまでのイントネーションパタンの研究では、句または文全体に渡ってピッチ周波数が徐々に降下する特性のフレーズ成分 (基底成分等とも言われる) 上に、アクセントを担う語や文節の音調成分 (アクセント成分) が重畳するモデルで記述されてきた。しかし、このようなイントネーションの記述は、朗読など「紙に書かれたテキストを読む」音声 (Read speech) の分析による影響があるのではないかと考えている。自発発話音声の句音調においては、こうした漸次的にピッチが下降する特性は必ずしも見られない (佐藤 2018A)。図 1 図 2 に、男性話者 : MA と女性話者 : FA の句音調の一例を示す。図中、母音、撥音部のピッチは 10 msec 毎の○印で、子音のうち接近音、鼻子音等はグレーの塗りつぶしで示した。発話内容を示す母音・子音記号中の記号 ] はアクセント記号であり、図 2 中の記号 [ はアクセントではなく句内の現れる上昇音調を示す。

二つの図が示すように、自発発話においてはアクセントのある音調区分ごとにピッチの上昇→アクセント核近傍でのピーク→安定な終結部に向けての下降、で示されるような (上昇-下降) パタンの連続するケースが多く観測された。朗読等の音声ではイントネーションを制御しつつ、複数のアクセント句から成る比較的長い区分をまとめて発話するのに対して、自発発話では、短い音声区分毎に発話が完成され、その区分の連続によって発話単位が構成されているように見受けられる。図 1 の話者 MA の場合、フィラー的発話「そう」と発話単位末の上昇境界音調の要素「ものね」を除くと、「みんなに」「夢を与

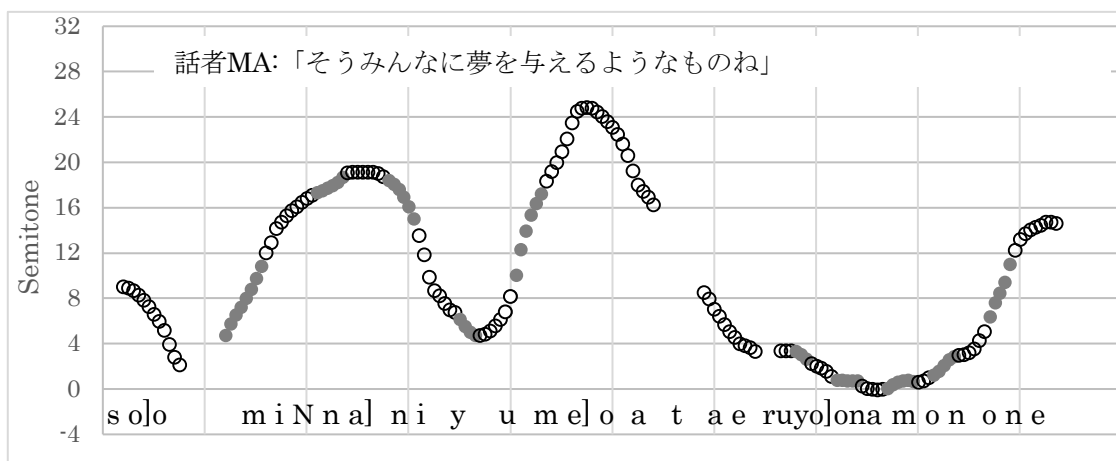


図1 句音調の分析例—卓立型音調様式の場合1 (話者：MA)

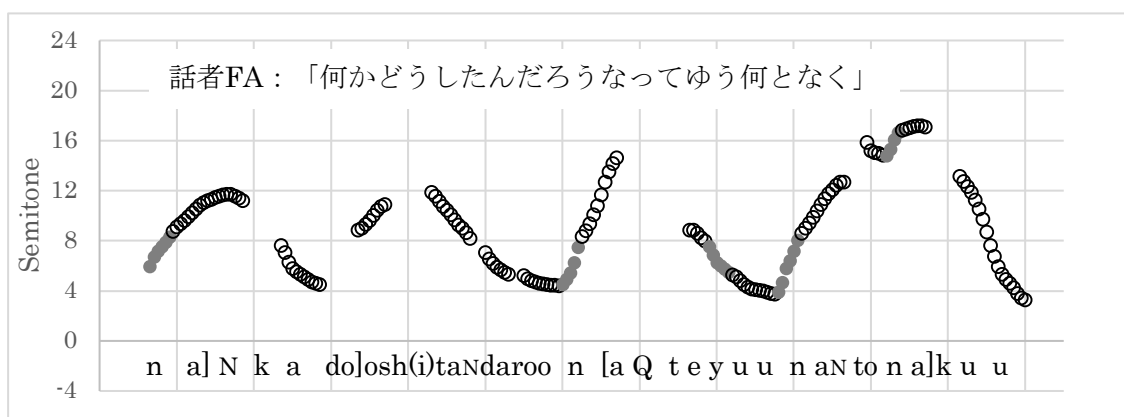


図2 句音調の分析例—卓立型音調様式の場合2 (話者：FA) ((i)は無性化母音を示す)

えるような」の二つのアクセント句から成り立っている。「みんなに」は、4 ST ほどの高さから上昇開始があり、開始と同じ4 ST の高さに戻る。「夢を与えるような」も同様に4 ST から再び上昇し、「与える」で4 ST に戻ったのち「ような」で終結音調の0 ST に降下したあと句末上昇に転ずる。図2の話者FAでは、話者MAと同様にアクセント句の上昇開始は4~6 ST 程度であり、上昇ののち4 ST ほどに降下するパタン連続となっている。話者MAのように0 ST ほどまで降下していないのは、まだ発話終結には至らず後続の発話単位につながっていくためと考えられる。このように大きなピッチの上昇と下降による際立ったアクセント句が連続する句音調特性を、ここでは卓立型音調様式と呼ぶことにする。自発発話における典型的な句音調の形式である。

卓立型音調様式では、句の始まりの高さは、強調の度合いや、アクセント句の句頭音調が低く始まるかどうかなどによって様々であるが、多くは4~6 ST 程度の高さで始まり、これより高い音調領域で(上昇—下降)の連続する句音調パタンが展開される。これより低い4 ST 以下の領域は発話単位を終結へ導く音調下降領域となっている。句が開始されて再び降下するおおよそ4 ST の高さは、発話努力の要しない楽に発声できるレベルであるが、発話単位や発話段落の終わりの0 ST 程度へと向かう音調下降では、声を低める一定の発話努力が必要なのではないかと考えている。ちなみに0 ST を音階の「ド」とすると4 ST の高さは「ミ」の音階に相当する。

一方、自発発話においては、卓立型音調様式とは異なりメリハリのない単調な発話様式も見られる。



図3 句音調の分析例—非卓立型音調様式の場合 (話者: MA)

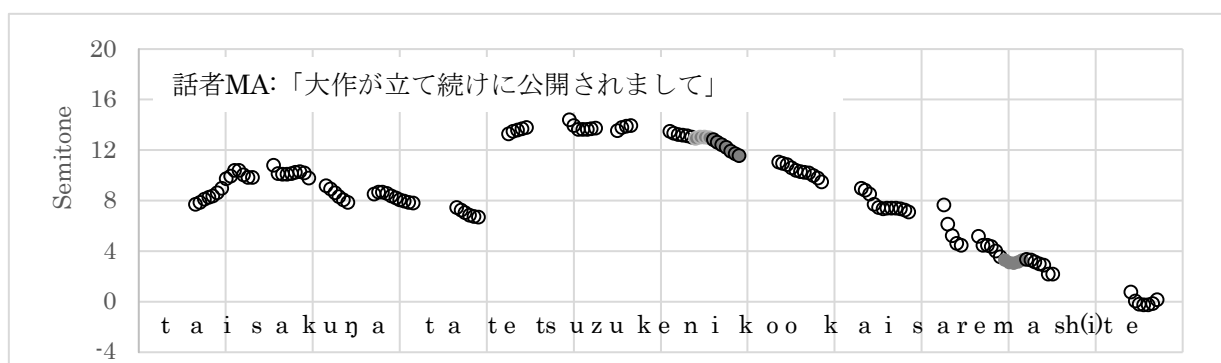


図4 平板型アクセントが連続する場合の句音調の例—漸次降下型音調様式 (話者: MA)

図3はその例であり (話者: MA), アクセント句を際立たせるような発話ではなく, 4ST程度を中心に, 音調の小さい上下を繰り返しつつピッチの終結音調に達する発話である. 印象としては, 独り言のような発話, 前置きの発話, ぼそぼそと呟くような発話であり, 主要句のあとに続く従属句の音調などにも見られる. このような句音調を非卓立型音調様式と呼ぶことにする. 卓立型, 非卓立型いずれの音調においても明確な漸次的下降特性は必ずしも認められない.

卓立型音調の場合, アクセント句の音調にはアクセント感を実現する下降と句音調の終結感に寄与する下降が見られる. 両者ははっきりと分離されているとは限らないが, アクセント核に伴う下降音調が契機となって, つまりアクセント核があってはじめて句や発話単位の終結下降も実現されている. それでは, アクセントがない句の発話の場合はどうなるであろうか? 実現例は少ないが図4がその一例である. 「大作が」と「立て続けに公開されました」の二つの句はともに平板型アクセントであり, アクセントに伴う音調の急峻な下降はない. そのため, 特に後者の句「立て続けに公開されました」では終結音調 0 ST に向けて漸次的に音調が降下する. このように明瞭にアクセント核が存在しない句の場合には, 従来からイントネーションモデルで示されてきたような終結音調に向けての右下がりに漸次的に降下する音調特性が観測される.

#### 4. アクセント拍の音調形式とアクセント位置

次に発話単位(UU)を構成するアクセント句の音調形式を見ていく. 例えば, 前節の図1におけるアクセント句「ゆめ]をあたえるよ]うな」では, まず2拍目にアクセントがあり, これは2型であってアクセント拍は上昇調で実現されているが, 「うな」は1型でありピッチ下降量は少ないが下降調で実現さ

れている。また同図中、3型である「みんな」にはアクセント拍内音調は平坦な音調で実現されている。このようにアクセントがあるとされる拍の音調は様々な形式で実現されることが予想される。アクセント感覚は、アクセント拍と後続拍との間の高さの相互関係で決まると考えられるが、まずアクセント拍内の音調の形式とアクセント位置との関係を調べた。具体的には、アクセント拍の拍内のピッチ周波数を直線近似し、以下の音調形式に分類した（佐藤 2018A）。

- ・下降音調(Falling Pitch: FP)：下降ピッチの音調形式
- ・平坦音調(Level Pitch: LP)：拍内ピッチの傾きの絶対値が 0.1 ST/フレーム 未満の場合を平坦のピッチとした。
- ・上昇音調(Rising Pitch: RP)：上昇ピッチの音調形式
- ・その他、上昇・下降音調(Rising+Falling Pitch: R&FP)、平坦・下降音調(Level+Falling Pitch: L&FP)なども設定したが、これらは数が少ない。

図5は1型（頭高）アクセントの「猫」という語を含む種々の句における音調パターンである（話者：FB）。アクセントのある第1拍目の音調形式は、左図の2例では下降音調(FP)と平坦音調(LP)であり、右図の2例はともに上昇音調(RP)で、そのうちの一つ「猫も」では後述するアクセントの遅下がり現象を示す音調パターンとなっている。すなわち、アクセントのある1拍目の/ne/よりも2拍目の/ko/のピッチの方が高く、見かけ上2拍目にアクセントがあるような特性となっている。このようにアクセントのある第1拍は同一語でも様々な音調形式で実現されている。なお、右図のようにアクセント拍の音調が上昇となる場合は、左図の下降音調や平坦音調の場合に比べてアクセントのプロミネンスが大きくなる傾向のあることを指摘しておきたい。

アクセント位置（いわゆるアクセント核の位置）は、アクセント拍内の音調形式によって著しく影響される。図6は図5と同一話者(FB)の場合の結果であり、1型アクセントにおけるアクセント位置がアクセント句のどの位置に生起するか、その生起度数を示したものである。

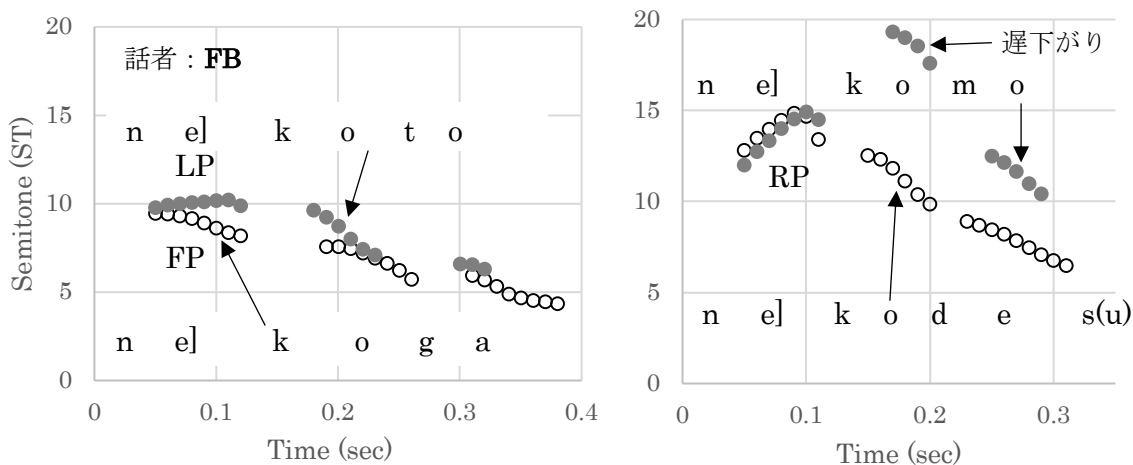


図5 アクセント拍の各種音調形式

（「猫」を含む各種アクセント句の音調、白丸とグレーの塗りつぶしのパターンで示す）

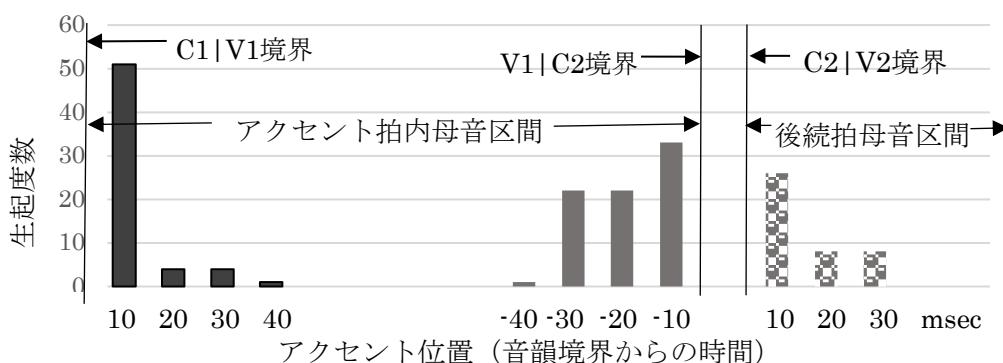


図6 アクセント位置の生起位置とその生起度数 (1型アクセントの場合 話者:FB)

横軸は音韻境界から測ったアクセント位置を 10 msec 毎のフレーム単位で示している。境界に隣接する直近フレームは±10 msec となる。アクセントのある第 1 拍を C1V1 (C:子音, V:母音), 後続拍を C2V2 とし, さらに C1|V1 をアクセント拍の子音と母音の拍内境界, V1|C2 をアクセント拍末の拍間境界すなわちアクセント核のある母音と後続子音の拍間境界, また C2|V2 を後続拍の子音と母音の拍内境界とする。図から明らかなように, アクセントはアクセント拍の中心付近に置かれるのではなく, アクセント拍母音の開始付近, 後続子音との境界付近, そしてアクセント拍に後続する拍の母音の開始付近に多く生起していることが分かる。

C1|V1 境界近傍にアクセント位置が分布する音声は, その 90% が下降音調であり, V1|C2 境界近傍にアクセント位置がある音声のうち, その 72% が上昇音調で 27% が平坦音調であった。C2|V2 境界の右側すなわち V2 に分布するアクセントは所謂「遅下がり」のものであり, その 72% は上昇音調であった。2 型アクセント, 3 型以上のアクセントにおいてもほぼ同様の結果が得られた。3 型以上のアクセント拍では, その殆どが下降と平坦の音調であり上昇調のものは少ないが, 後接要素が付いた複合形の場合, その要素が強調されると上昇調となる (例:「思いま]して」, 「そんなこ]と」における「ま]して」「こ]と」が強調発声された場合など)。

以上示したように, アクセント拍の音調形に関して, それが下降調の場合, アクセント位置は当該拍の母音の開始点近傍に分布し, 上昇調の場合は当該拍の母音の末尾近傍に分布するか, もしくは後続拍の母音部にまで達する。平坦調の場合は, 当該拍末尾近傍にアクセントのくる場合が比較的多いことが明らかになった。特に平坦調や上昇調の場合, 後続拍内のピッチ下降がアクセント感の生成に必要であると考えられ, アクセントの音調単位は後続下降拍を含む 2 拍のまとまりで考える必要があるのではないかとと思われる。またピッチの下降時点が音韻境界近傍に集中することは, 音調変化の制御と音韻に関わる調音制御が同期的に行われていることを窺わせる。これらについては後続の諸節でまた触れることにする。

## 5. アクセントの「遅下がり」現象の生起傾向と生起要因

前節でアクセントの「遅下がり」現象の一例を示したが, 本節ではこの現象に焦点を当ててその生起傾向と生起の要因に関して述べる (佐藤 2018B)。

表 1 は発話者:FB の場合の遅下がりの生起数をアクセント型毎に示したものである。この発話者の場合, 合計 81 例の遅下がり確認されたが, これは発話の全アクセント単位の 15.6% に当たる。なお, アクセント型は先頭から数えたアクセント拍の位置で示した (1 型ア, など)。表 1 から「遅下がり」は 1 型アクセントで著しく多く (72.8%), また 1 型アクセントの全抽出例のうちその 1/4 が「遅下がり」であった。2 型アクセントがこれに続き多く, 3 型以上ではその生起数は減少する傾向がある。

表1 アクセントの「遅下がり」生起数（アクセント型別，話者：FB）

	1型ア	2型ア	3型ア以上	計
遅下がり生起数	59 (72.8%)	11 (13.6%)	11 (13.6%)	81

表2 アクセント拍の音調形と「遅下がり」の生起数の関係（話者：FB）

アクセント拍音調形	1型アクセント			2型アクセント		
	F	L	R & RF	F	L	R & RF
遅下がり生起数	3	7	47	0	2	9
割合	5.1%	11.9%	79.7%	0%	18.2%	81.8%

表2は、同一話者におけるアクセント拍での音調形と遅下がり数との関係をまとめたものである。ここで(R&RF)の欄は上昇に関わる二つの音調（上昇および上昇・下降）の生起数である。1型および2型アクセントとも、上昇に関わる音調(R&RF)の生起数が全体の約8割を占め、アクセント拍での上昇関連音調と遅下がり現象との間に強い相関のあることが分かる。つまり、アクセントの遅下がりとは、アクセントが1型か2型であって、かつアクセント拍が上昇音調に関係する場合により多く生起するということになる。1型アクセントでは句頭の最初の拍で「高」のトーンが要請され、2型アクセントでは1拍目が「低」く始まるが、後続のアクセント拍で「高」音調を実現しなければならない。両者とも句の最初に現れるアクセント拍が「高」であり、この「高」を上昇音調で達成することと「遅下がり」が関わっていると考えることができる。次にアクセントの「遅下がり」現象がどのような条件で生起するか、以下具体的にその諸要因に関して述べていく。

### (1) 調音制御と遅下がりの関係

はじめに遅下がりに関わる音調制御と音韻生成に関わる調音制御との関連を見ていく。図7は女性話者FAの発話例である。発話段落は三つのアクセント句「かなり」「その検査が」「長くて」から成っているが、発話のまとまり（韻律句）としては「かなりその」「検査が」「長くて」のような音調区分で発話されている。アクセントのある語は1型であるが、いずれもアクセント拍の後続拍にアクセント位置が観測される「遅下がり」の現象を示す。図中↑印はアクセント位置を、点線の矢印はアクセントのある語頭拍と後続拍におけるピッチ変化を示しており、いずれも上昇調ののち下降する。「かなり」では、音調がアクセント拍の語頭母音/aから後続子音/nに至るまで上昇しており、第2音節母音に至って音調降下の契機（アクセント位置）が認められる。同様にして「検査が」では第2拍の撥音/N上に、「長くて」

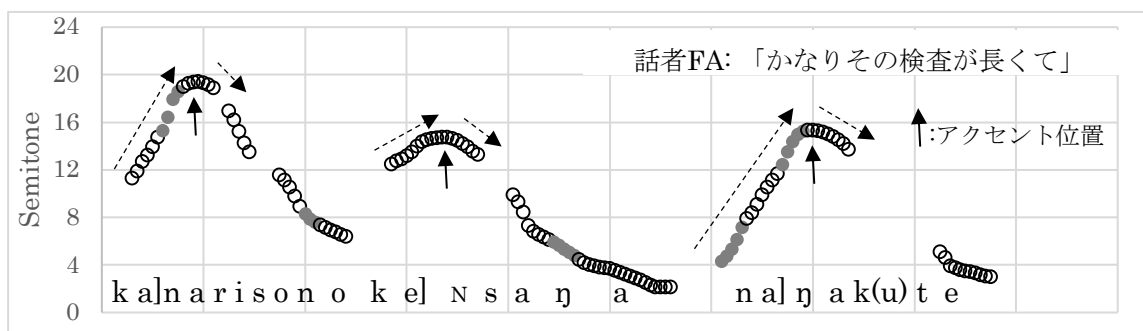


図7 1型アクセントにおける音調の上昇・下降パターンと「遅下がり」



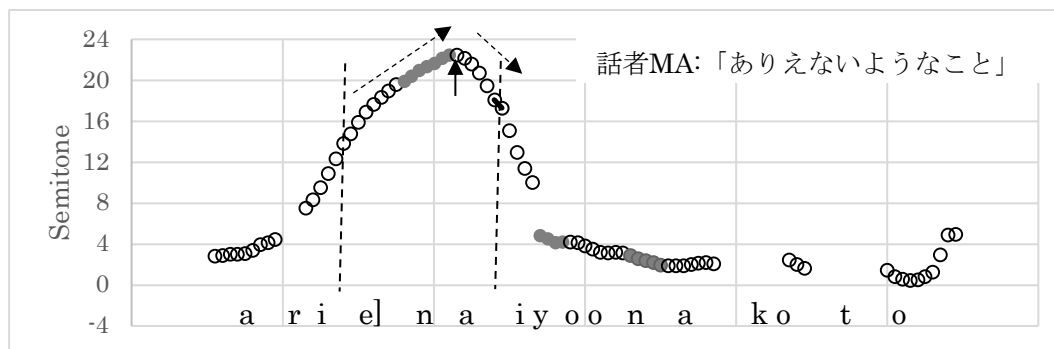


図8 3型アクセントの発話における「遅下がり」

では[na]jaの第2音節の母音開始部にアクセント位置がくる。つまり音調が音節の開始点ではなく音節内母音や撥音の onset 近傍に同期して制御され、加えて下降音調に向かう音調の変曲点がアクセント拍内ではなく後続母音内（あるいは撥音内）に現れることが多く、それが「遅下がり」をもたらす。日本語の音節の基本は（子音－母音）の CV 構造であるが、発話のタイミングは音節の頭である子音の開始点というよりは、後続母音の開始を契機としていることと関係していると考えられる。

以上、1型アクセントの発話における「遅下がり」をみてきたが、それ以外のアクセント型でも同様の現象が見られる。図8は、男性話者 MA による3型アクセントの語を含む例「ありえないようなこと」の音調パターンである。/arie]nai/における3拍目の/e/にアクセントがあるが、/e/ではアクセントに伴う音調が上昇し、その上昇が後続子音の/n/まで持続する。ピッチが反転下降する変曲点（アクセント位置）は後続母音/a/の onset 近傍でありアクセントの「遅下がり」になることが分かる。

以上二つのアクセント型の発話データは、アクセントの「遅下がり」現象がアクセント実現のための下降音調制御と母音音高が同期的に制御されていることと密接に関連しているのではないかということを示している。

## (2) アクセント拍の上昇調との関係

(1)において「遅下がり」現象を音調制御と調音制御の観点から記述したが、これと関連してアクセント拍における上昇音調との関係を述べる。特に「遅下がり」は1型アクセントで多く、さらにアクセント拍が上昇調の場合に多いことはすでに触れた。1型アクセントでかつ上昇拍の場合の例を図9に示す。図中2例とも同一話者 FB の「我が家の」という発話である。(a)は卓立型発話でプロミネンスが大きく、「遅下がり」が観測される。(b)はプロミネンスが小さい非卓立型発話であり、この場合「遅下がり」が見られない。アクセント拍が同じ上昇調であっても、プロミネンスが小さい場合、アクセント位置は当該拍末近傍に止まるが、アクセント位置が後続拍に移動する「遅下がり」が見られる場合は一定以上の急峻な上昇音調を伴うと考えられる。この場合「遅下がり」は音調上昇制御の overshoot 現象と捉えることも可能である。

1型アクセントが上昇調となるのは、句頭の最初の拍における「高」の音調感を上昇音調で作りに出しているからであろう。2型アクセントの場合は、1拍目は「低」、2拍目で「高」のトーンが要請されるため、1型アクセントと同様にアクセント拍が上昇調で実現され「遅下がり」が観測される。なお、2型では1拍目から緩やかな上昇が見られることがあり、1型アクセントほど急峻な特性にはならないことがある。

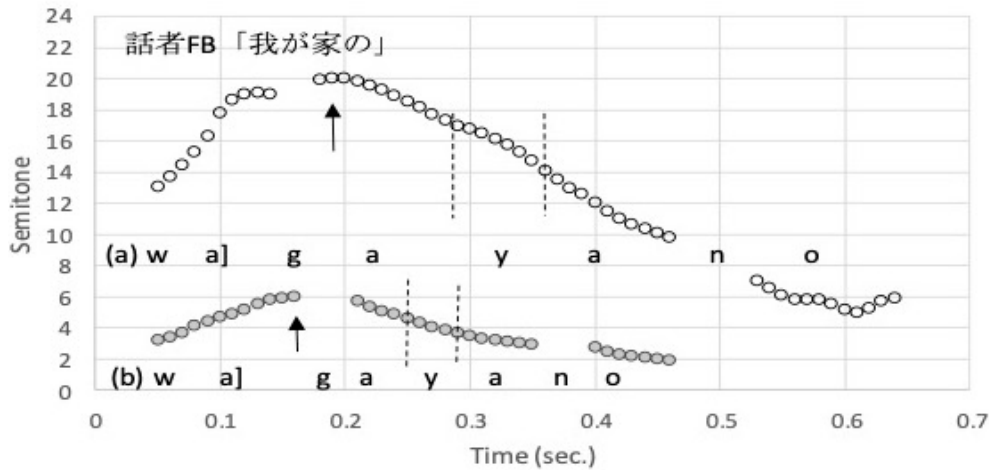


図9 1型アクセントの語頭上昇音調と「遅下がり」

1拍目で上昇調となる1型アクセント語のアクセント知覚に関する実験では、1拍目で十分急峻な音調の上昇がある場合、後続拍の許容される音調下降には大きな幅があり、わずかな下降でも1型アクセントの知覚が得られること、また比較的緩やかな上昇特性の場合、後続拍に一定以上の下降特性が求められることが分かっている(佐藤 2016, 2017). このことは、1拍目の急峻な音調の上昇があれば、それがアクセント知覚の Cue として役立つので、後続拍が「遅下がり」に伴う音調下降は緩やかであってもかまわないということを示している. 従来から、アクセント感覚に関しては音調の「下降」に着目して論じられてきたが、「上昇」特性のアクセント感覚生成に果たす役割についても今後更に目を向ける必要があるものと思われる.

### (3) フット単位の発話との関係

東京方言アクセントの特徴として、長音、撥音等を含む長音節にアクセントがある場合は、これらの従属拍にアクセントが置かれることはなく、前部拍がアクセント拍となる. しかし、実際の発話の際には、この2拍の音節は一体となって発音されるため、音響特性から見たアクセントはしばしば後部従属拍に位置することがある. これもひとつの「遅下がり」と考えられる. 図10は、2型アクセントの発話「デパートの」の音調パターンである. 音調は第2拍の母音 onset から長音部まで上昇を続け、アクセント位置は長音節/pa] a/の音節末近傍にある. 音節内では音調下

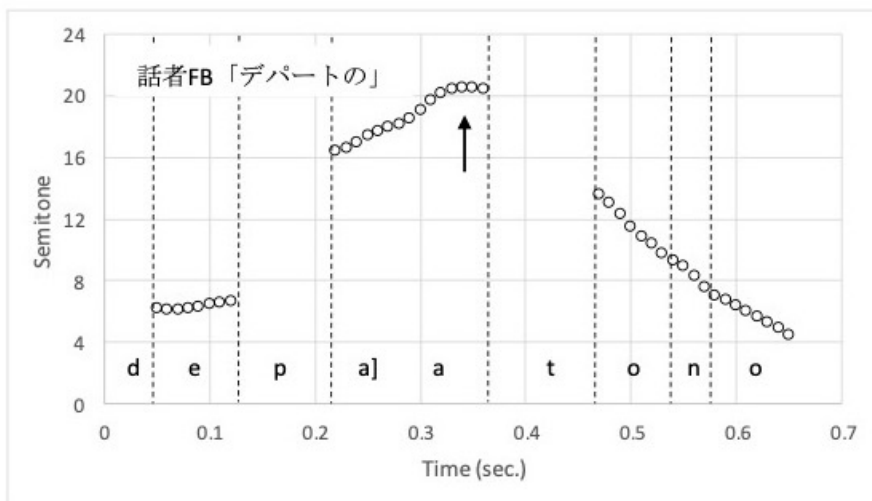


図10 長音節にアクセントがある場合の「遅下がり」例

降である. 音節内では音調下

降はほとんどなく, 急峻な音調降下はアクセント拍より 2 拍遅れた拍/to/に見られる. このようにアクセントのある 2 拍の長音節が上昇調で発話される場合には, 「遅下がり」的現象がしばしば観測される.

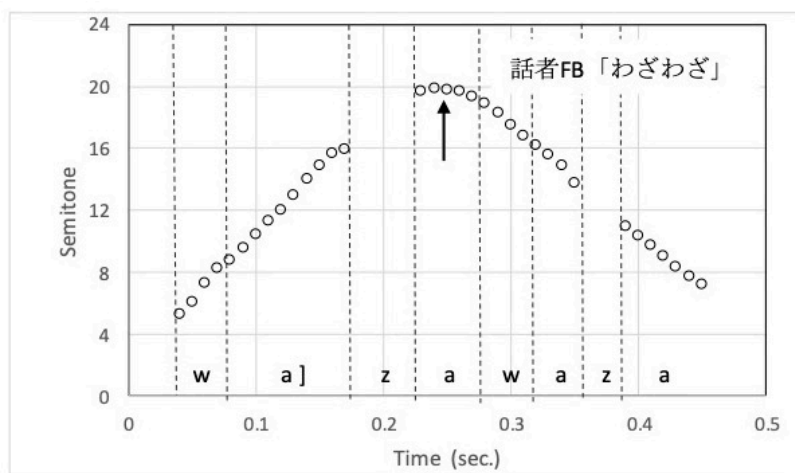


図 11 フットリズムの発話と「遅下がり」(1)

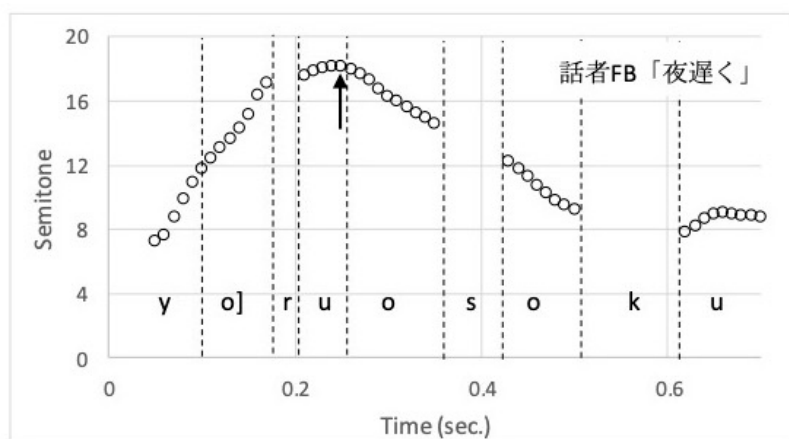


図 12 フットリズムの発話と「遅下がり」(2)

上記で述べた従属拍を含む長音節での「遅下がり」ばかりでなく, 通常の単音節が 2 拍ずつまとまって発話される場合にも同様の現象が見られる. 日本語のリズムは, 拍の等時性に基づく拍リズムがその基底にあると考えられるが, 実際には 2 拍まとまってリズム単位となる発話がなされやすい. これがフット(脚)リズムである. 図 11 と図 12 にフットリズム発話の場合の音調と「遅下がり」の例を示す.

図 11 は「わざわざ」と発話された例であり, (わ)ざ(わ)ざのような 2 拍毎のフット単位で発話されている. 音調は最初のフット/wa]za/内で上昇を続け, 2 拍目の/za/の途中で下降の契機を示すものの下降量は少なく, 音調下降は後続フットの/waza/で生ずる. アクセント位置は 2 拍目の/za/にあり「遅下がり」である. 図 12 は「夜遅く」の発話例であるが, (よ)る(おそ)(く)のようなフットリズムで発話されて

いる. 句頭のフット/yo]ru/では, アクセント拍の/yo/で急峻な音調上昇があり, 後続拍/ru/においても緩やかな上昇が続き音調下降はない. 下降は後続するフット/oso/内で生ずる. これら二つの発話例から, アクセントに伴う音調の上昇→下降の制御は, 拍単位というよりも 2 拍まとまったフット単位で行われることが原因で「遅下がり」という現象をもたらしているのではないかと考えられる.

#### (4) 強調的発話との関係

(1) 項で卓立型発話の場合に上昇音調アクセントで「遅下がり」が生じやすいことを述べた. 卓立型は句の音調パターンを際立つように発話する様式であるが, 特定の句を意図的あるいは対比的にいっそう強調して発話する場合, アクセントの大きさ(プロミネンス)はさらに高まる. アクセント句のピークの高さが 20 ST を超えると強調の印象が強く感じられる. 音調の上昇幅や下降幅が 1 オクターブまたはそれ以上(12 ST 以上)になるとその動態特性の影響は後続拍にまで及び, ほとんどの場合「遅下がり」現象を呈した. 一例として, 「私の枕を使って」における「枕を」の部分で強調的に発話した場合を図 13 に示す.

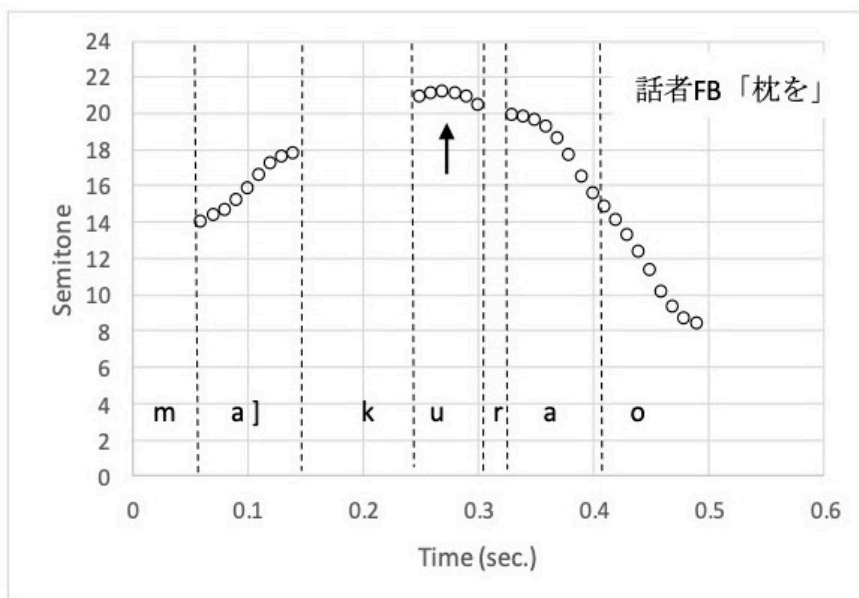


図13 強調的発話における「遅下がり」例

アクセント拍の/ma/は上昇調で発話され、後続拍/ku/でいっそう高いプロミネンス値となり、アクセントの「遅下がり」現象を示す。

既に例示した図1における発話における「夢を与える」の部分、および図8における「ありえないようなこと」等もプロミネンス度が22 ST程度あり、強調的発話として「遅下がり」が観測される例である。

(5) アクセント拍での音調上昇がない場合の「遅下がり」

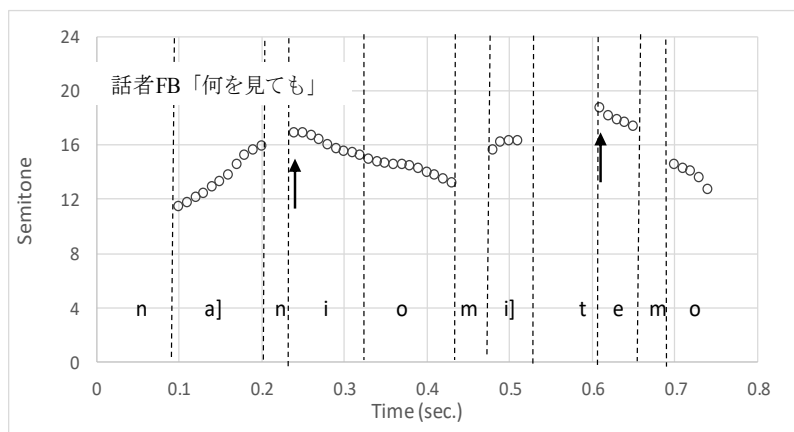


図14 アクセント拍が上昇調でない場合の「遅下がり」(1)

これまで示してきた「遅下がり」の実例では、アクセント拍が上昇音調を示す場合のみであった。表2で示されているように、80%は上昇調であるが、下降調(FP)、平坦調(LP)の場合も存在する。ここでは、上昇調以外の場合に「遅下がり」現象が生起する実例を見ていく。図14は、「何を見ても」の例であり、音声区分「何を」と「見ても」のいずれも「遅下がり」である。

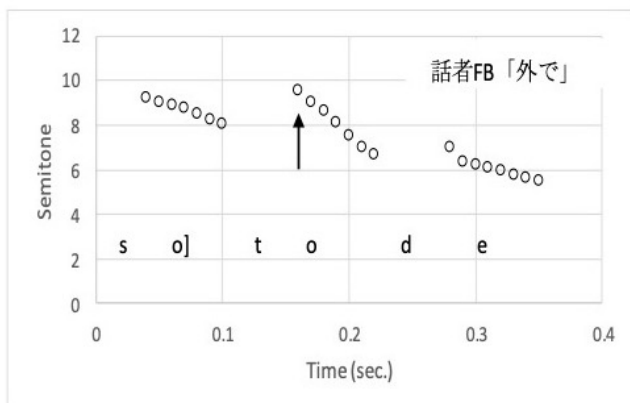


図15 アクセント拍が上昇調でない場合の「遅下がり」(2)

後者の「見ても」では、アクセントのある/mi]/の母音部で音調上昇がないため、後続拍においてアクセント生成のため高いピッチからの降下が必要となる。これが後続拍頭にアクセント位置が見られる原因となっている。アクセント拍が上昇調のときは、後続拍で顕著なピッチ降下の認められない場合もあるが、明瞭な上昇調がない場合は後続拍での音調下降特性が必要となる例である。

図15は、非卓立型発話の「外で」の例である。この場合は、アクセントに後続する母音onsetのピッチが高くなって「遅下がり」的現象を呈する。これは、無声閉鎖音では開放後の

母音のピッチが高まるという *micro prosody* の事例となっているが、後続拍でのこのような高いピッチからの音調降下はアクセント感の生成にも寄与しているであろう。

アクセント拍自体が下降音調の場合は、その下降がアクセント感の生成に役割を果たしていると考えられるが、ここで示したようなアクセント拍が緩やかな下降調や平坦調の場合には、後続拍における音調降下がアクセント生成のために必須となると考えられる。

## 6. アクセント位置以降の下降音調

すでに触れたように、日本語アクセントの型は、拍の「高」と「低」のトーンの配置、あるいは「高」から「低」に遷移する際の「高」の位置（アクセント核）等によって表現されるが、いずれにせよ「高」の拍から「低」の拍への音調降下が語のアクセントと深く結びついていると考えられている。それでは、いったいどの程度の音調の降下がアクセント実現のために必要なのであろうか。ここではアクセント位置以降の音調の下降について見ていきたい。

前節で述べた「遅下がり」現象の一例である図7を見てみると、音調パターン上のアクセント位置から次の拍末への音調降下量は3つのアクセント句とも2 ST 以下であり非常に小さい。図16と図17は、同じ話者FAの発話例であり、アクセント位置と後続拍の音調降下の例を示したものであるが、同様の性質が見られる（↑：アクセント位置，↓：後続拍末）。

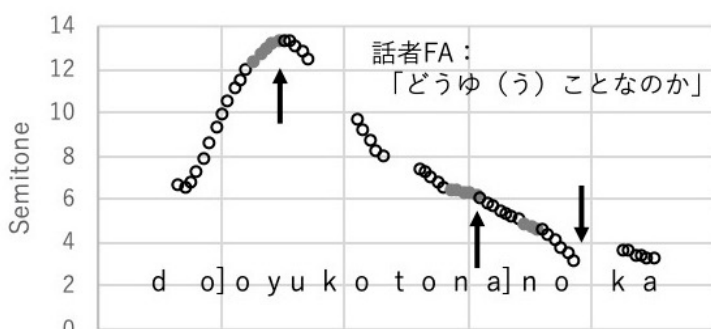


図16 アクセント位置と後続拍への音調降下例(1)

図16の最初のアクセント句「どうゆ(う) /do]oyu/ は1型アクセントであるが、最初の2拍の長音節/do]o/は一体となって上昇音調で発音され、音節末まで上昇が続く。第3拍の母音/u/に至って音調降下が生じる。その降下量は2 ST 以下である。

図17では最初のアクセント句「お箸が /oha] shiŋa/ もまた「遅下がり」であり、アクセント位置の母音内音調降下は3 ST 程度であるが、アクセント句末までは11 ST 程度と大きく降下している。

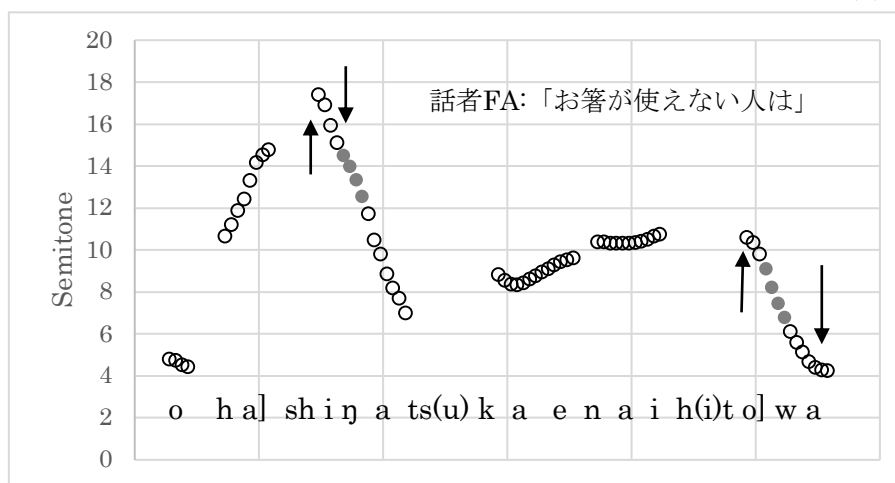


図17 アクセント位置と後続拍への音調降下例(2)

次のアクセント句「人は /h(i)to] wa/ では、アクセント拍/to/の拍内音調は下降音調でその降下量は小さく、後続拍/wa/末までの降下量は7 ST 程度と大きい。

以上のように、音調上のアクセント位置から後続拍への音調降下の大きさは様々である。図17の「人は」において音調降下が大きいのは、ア

アクセント感覚生成のための音調下降と句末の終結のための音調下降が一体となったためであると考えられる。アクセントの役割は (a) 語の弁別, (b) 語や句に纏まりを与える, の二つがあると考えられているが, このうち後者の役割は, 単にアクセントそのものが纏まりという機能を果たしているのではなく, 実際には音調パターン上でアクセント句としての「纏まり感」や句の「終結感」をもたらす音調下降特性, もしくは終結値への到達があつてはじめて実現される機能であると考えられる。従つてアクセント核以降の音調降下には, アクセント弁別に直接関わる降下とアクセント句として終結させる降下とがあり, 両者を区別して考える必要がある。その上で前述の「人は」で見られた例のように, アクセント位置以降の音調降下を両音調降下の融合として捉えるべき例も多い。

アクセント位置からの後続拍末への音調の下降量 (拍間下降量) は, アクセント拍が上昇音調あるいは平坦音調の場合 2~3 ST であり, 下降音調の場合のアクセント拍内下降量は 1.4 ST 程度であることを報告した (佐藤 2018A)。図 16, 図 17 の例を含めて, このようなアクセントを生成する音調下降の量は, アクセント句の終結下降量と比べて小さい。アクセント感覚には, 音調下降量よりも音調パターンが下降に向かう契機 (音調変曲点) の役割が重要なのではないかと考えられる。音調下降に関して僅かな変化の契機しか示さない例ばかりでなく, 全く示さない例もある。これらはアクセント拍における音調の上昇調がアクセント感覚形成に一定の役割を果たしているとも考えられ, 今後のアクセント知覚研究の面からも更なる検討が必要である。

## 7. アクセント句の句頭音調

これまで主にアクセント位置を音調パターンの変曲点との関わりから述べてきた。音調の変曲点は, アクセント句が複数連続する場合, 句頭の音調上昇や句末の終結下降に関わる現象としても観測される。図 18 は「気持ちはよく分かる」 (話者 MB) という発話の例である。/kimochiwa/ から/yo]ku/へ変わるところで, /yo]ku/の/y/からピッチが上昇するのではなく, /y/は先行発話の終結下降の一部として融合下降し, 次のアクセント核の/o/で上昇に転ずる。この場合も母音 onset と音調制御時点が同期している (図中, 下降-上昇の点線矢印で示す)。さらに, /yo]ku/から/waka]ru/への連続では, 後者の第 1 拍の/wa/が先行する/yo]ku/と融合し, そのアクセント句の終結下降の役割を担っていると同時に, その下降が/waka]ru/の第 1 拍目の「低」の音調感の実現にも寄与していると思われる (図では下降矢印で示した)。

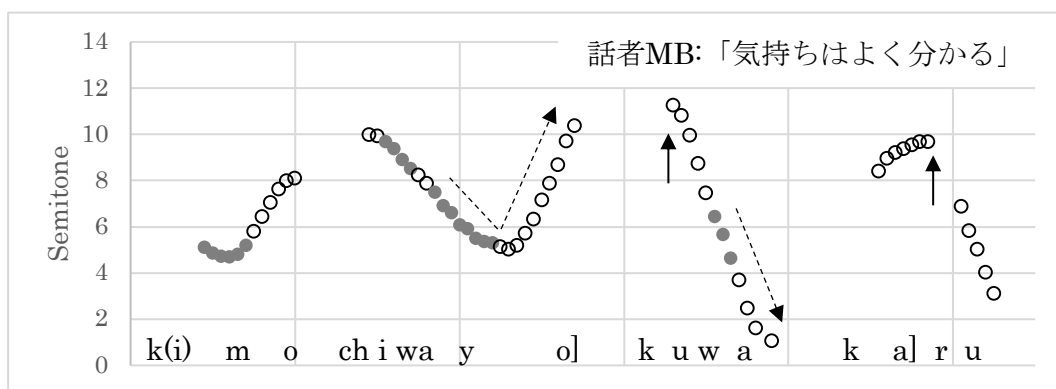


図 18 アクセント句連続の場合の第 1 拍目の音調特性例

図 19 は「みんなに」「夢を」「与えるような」という発話の例であり, 図 1 の発話の主要部分を再掲したものである。この中で「夢を」が主要なアクセント単位であり, 「与えるような」は従属的な要素である。/me]o/における上昇・下降音調に続いて/ataeru/が下降音調として接続し終結下降の役割を果たしてい

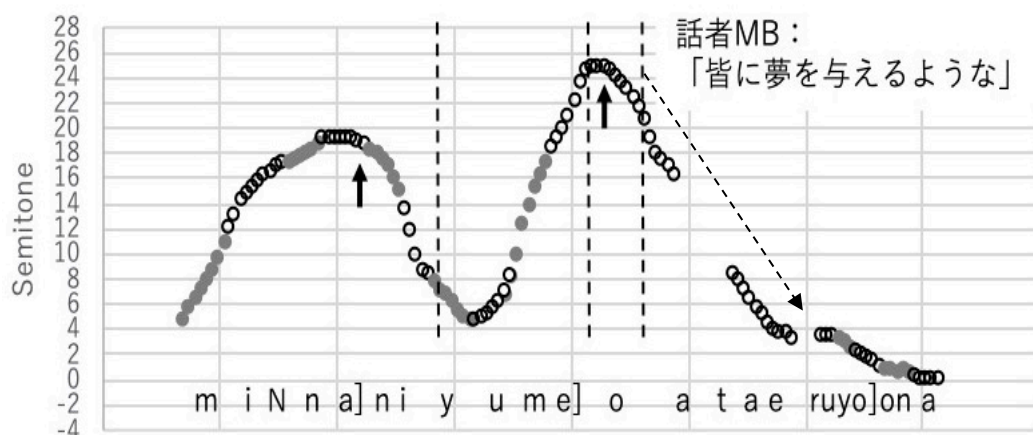


図 19 主要アクセント句に従属する句の音調パターン例 (図 1 の一部再掲)

る (点線下降矢印で示す). /ataeru/は平板アクセントであり, 1 拍目/a/は「低」であるが, 実際には 22 ST と基底値の 2 オクターブ程度の高い音調で接続している. このように, 1 型アクセント以外の語の語頭の「低」は, 丁寧な発音あるいは句頭等の場合にはその「低」が音響的に実現されるが, 自発発話のような連続音声では先行句の音調に依存して様々な高さとなり, 従属的, 不定的なトーンとしての振る舞いを示す (佐藤 2019).

なお, 同図で最初のアクセント句/miNna]ni/と次の/yume]oataeru/の境界における後続句頭拍/yo/の音調では, 図 16 の例と同じように子音/y/は先行音調に融合して下降音調となり, 母音/u/からアクセント句の上昇音調に転じている. アクセント生成に関わる下降音調の場合ばかりでなく, 音調が「高」へと上昇に転ずる場合においても音調制御と母音 onset の同期が見られることを指摘しておきたい.

## 8. アクセント句の構造モデル

アクセント句の音調を調べるうえで, アクセント句の構造を想定し, それを手掛かりに音調パターンを見ていくことが有効ではないかと考える. これまで述べてきたアクセント句の音調の諸特性の結果をまとめて記述するうえでも有用であろう. アクセント句を想定するポイントは以下である.

- ・アクセントの生成は (上昇-下降), (平坦-下降) など 2 拍に渡る高さの変化によるとの考えから, アクセントが生成される単位は原則 2 拍からなる単位とし, これをアクセントの動態核と呼ぶ.
- ・動態核に続く音調下降を二種類とし, アクセント単位の音調上のまとまりに関わる音調下降とこれを含む句としての終結感を与える音調下降とに分ける.
- ・2 型以上のアクセントの場合上昇調を伴うのが一般的であるが, 1 型アクセントの場合, 特に卓立的発話では句頭が上昇調で開始されることが多く観測されるので, 句頭に音調の上昇部分を設定する.

上記を配慮すると, アクセント句の基本的構造は以下ようになる (佐藤 2019). これは, アクセント音調と句音調が融合した形式である. 図 20 に, 実際の音調パターン上でのアクセント句の区分化 (構造) を示しているが, アクセント句が一定の卓立性を有する場合の音調形式を想定している.

開始(Onset)－上昇部(Rising)－動態核(Dynamic kernel)－  
 アクセント終結部(Accent coda)－句終結部(Phrase ending)

アクセントの動態核は、現在のところ原則 2 拍と考えておく。その 2 拍の動態核の音調パターンは、(a) 下降-下降 (b) 平坦-下降 (c) 上昇-下降 (d) 上昇-上昇、等のパターンが観測される。(a) と (b) は、非卓立型様式の発話、3 型以上のアクセントの語、平板型アクセント要素に後続する語などに見られることが多い。(c) と (d) は句頭の 1 型・2 型アクセントに見られるほか、句中で特にアクセントが強調された要素で観測される。アクセントを 2 拍間の音調変化を包含する動態単位で捉えることにより、拍単位のアクセント把握からくる「遅下がり」現象の概念は、特異な音声現象としてではなく、発話様式の自然な形式として理解がしやすくなるのではないかと考える。

またアクセント句構造化のパターンは典型的形式であって、どのような発話においてもこの構造モデルのまま実現するわけではない。例えば 2 型アクセントの場合、語頭の onset としての「低」の拍が上昇調で始まり、アクセント動態核の前部の上昇調と一体化することがある。また尾高型アクセントのようにアクセント拍の後続拍がない場合などで動態核が 2 拍とにならないことがある。さらにアクセント終結や句終結を担う要素がない場合、動態核の後部要素の下降音調の終結機能をも担わなければならないことがあり、この場合アクセント位置からの大きな音調下降がある。アクセント単位の終結下降と句の終結下降は必ずしも分離せず、一体化して両者の機能を担っていると思われる場合も多く観測される。句の音節構造、後部につく従属的要素の有無などによって現れる音調実態は様々であるが、ここで提案したアクセント構造を手掛かりにすると日本語アクセント理解に有用であろうと思われる。

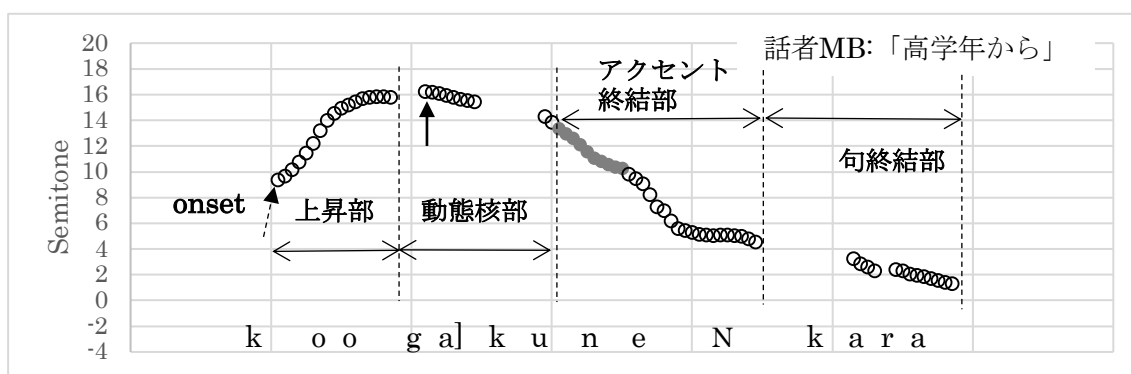


図 20 アクセント句の構造モデルとその音調パターン

## 9. アクセント句のその他の音調特性

最後に今回の分析で観測された音調特性における他の事象について紹介したい (佐藤 2018A)。音声を聞き取るとアクセントの可能性が二つあり、どちらか一つに決めかねた例である。アクセント句でアクセントが複数ある場合は、通常先行アクセントにおいて「高」から「低」への音調下降がいったんあり、そのあと後続アクセントへと続くためそれぞれのアクセントが認められるが、ここでの例はアクセント拍が「低」のトーンを介さず連続して連なる場合である。

図 21 に、「先日のことなんです」の音調パターンを示す。ここで、まず「ことなんです」は、/koto]naNdes/ のように/to/にアクセントがあるように聞こえるが、/kotona]Ndesu/のように/na/にアクセントがあるようにも聞こえる。筆者にとっては二つのアクセントが隣り合うため、同時に二つが知覚されることはなく、どちらか一方のアクセントとして聞こえる。従って一つのアクセントを定めることができない音調パターンである。これを仮に継起ダブルアクセント(Sequential double accents)と呼んでおく。

音響特性を見ると/to/から/na/にかけて一定の急峻な下降 (傾き:  $-0.44$  ST/FR, FR: フレーム 10 msec) があり、かつ撥音/N部ではさらに急峻な下降となる。このため、二つのうちのどちらのアクセントとしても判断することが可能となると考えられる。



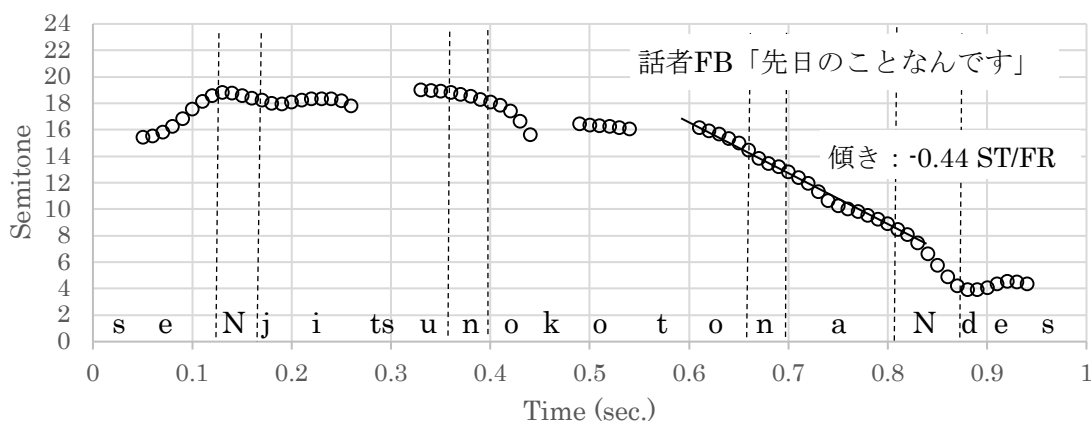


図 21 継起ダブルアクセントとその音調パターン

「～なんです」という後接要素は単独発話では「な」にアクセントがあるが、ゆっくりとした丁寧な発音では、前部に平板型アクセントの語がくる場合に「なんです」のアクセントが実現する（鼻+なんです→/hanana]Ndesu/）。一方、アクセントのある語がくると前部要素のアクセントが優先され、後部要素のアクセントは抑圧されてしまう（花+なんです→/hana]naNdesu/）。連続発声や自由発話になると、必ずしも抑圧されずにもともとあるアクセント核の性質が顕在化するため生ずる現象ではないかと推測される。

## 10. 議論とまとめ

自発的に発話された独話音声进行分析資料として、発話単位の句音調と句音調を構成するアクセント音調の諸特性に関して報告した。句音調に関しては、読み上げ音声とは異なる自発発話音声特有の二様の発話様式、すなわち卓立型と非卓立型の発話様式があることを示した。アクセント音調に関しては、連続発話音声における日本語アクセントの実現形式を音調特性の面から検討した。アクセント拍における音調形式を分析するとともに、アクセント位置がアクセント拍より後部の拍に見いだされる所謂「遅下がり」現象の生起条件を分析し、「遅下がり」は、発話の音調が母音調音の onset と同期的に制御されること、アクセント音調実現のためのピッチの上昇特性、2拍を1纏りとするフット単位の発話、ならびに無声子音に伴う母音の micro prosody 効果、等々と密接に関連することを明らかにした。これらの結果からアクセント句の音調構造の要素として、アクセントを生成する単位を2拍の動態核として捉えるモデルを提案した。しかし、動態核内前部の上昇調および動態核全体に渡る上昇調がなぜアクセント感覚をもたらすのかは未だ未解決の問題であり、アクセントの知覚実験などを含めた更なる研究が必要である。また、「遅下がり」は非常にゆっくりとした丁寧な発音では起こらないと考えられるので、「遅下がり」現象の発話速度との関連も今後の課題である。

アクセント句音調に関しては、上記の外、アクセントの動態核以降の音調の降下特性に関してアクセント単位のまとめり・終結に関わる音調下降と句終結に向かう音調下降とを区別する必要性のあることを論じた。また、日本語の1型アクセント以外アクセントは語頭のトーンは「低」で記述されるが、この語頭トーンは連続音声中では先行句に融合されるなど *dangling tone* として振る舞う性質があること、アクセント核が直接隣接して存在するアクセント句では、二つのアクセント型のうちどちらの型も認めることが可能である音調特性が存在することなどを報告した。

また、これらの日本語アクセント句の音調分析の結果から、アクセント動態核を含む日本語アクセン

ト句における音調パタンの構造的モデルを提案した。これによりアクセント位置（核）の移動，アクセント生成における上昇調の役割，アクセント核以降の音調降下の役割区分などにおいて，理解のし易さと検討の枠組みを提供したと考えている。

今後更に分析例を増やすとともに，統計的データの集積，音声合成手法を援用したアクセント知覚実験などによって日本語のアクセントを音調の動的振舞の面から検討を進める予定である。

**謝辞** 本研究は，科学研究費助成金（基盤 C）「音調動態形式に基づく日本語アクセントの研究」（19K00598）の助成を受けたものである。また，研究にあたって国立国語研究所より「日本語話し言葉コーパス」の提供を受けるとともに，同研究所のコーパス開発センター共同研究プロジェクト「コーパスアノテーションの拡張・統合・自動化に関する基礎研究」の一環としても実施したものである。お世話になった関係各位に厚くお礼申し上げます。

#### 参考文献

- 五十嵐陽介・菊池英明・前川喜久雄(2006). 「報告書 日本語話し言葉コーパスの構築法 『第7章 韻律情報』」, URL: [pj.ninjal.ac.jp/corpus\\_center/csj/document.html](http://pj.ninjal.ac.jp/corpus_center/csj/document.html)
- 佐藤大和(2016). 「共通語における動的音調形式とアクセント知覚」, 日本音声学会 第334回研究例会, 2016.12, 於: 十文字学園女子大学
- 佐藤大和(2017). 「アクセント核のあとピッチの急峻な降下はあるか? —ピッチの動態特性とアクセント知覚—」, 3-8-4, 日本音響学会 2017 春季研究発表会講演論文集
- 佐藤大和(2018A). 「アクセント音調の諸相とその動態形式」, 「言語資源活用ワークショップ 2018」論文集, 国立国語研究所, pp.592-599
- 佐藤大和(2018B). 「自発発話データから見たアクセントの遅下がり現象」, 第32回日本音声学会全国大会予稿集, pp.66-71
- 佐藤大和(2019). 「自然発話におけるアクセント音調の動態分析」, 「言語資源活用ワークショップ 2019」論文集, 国立国語研究所, pp.30-37
- 杉藤美代子(1980). 「“おそ下がり”考 —動態測定による日本語アクセントの研究」, pp.201-229, 徳川宗賢編「論集日本語研究2 アクセント」, 有精堂出版

執筆者連絡先: [sato.hirokazu@tufs.ac.jp](mailto:sato.hirokazu@tufs.ac.jp)

原稿受理: 2020年12月14日