



## Kobe Shoin Women's University Repository

Title	音声自然降下の知覚的補正に関する予備的研究 A Study of Perceptual Revision of Declination
Author(s)	松井 理直 (Michinao Matsui)
Citation	Theoretical and applied linguistics at Kobe Shoin, No.4 : 77-84
Issue Date	2001
Resource Type	Bulletin Paper / 紀要論文
Resource Version	
URL	
Right	
Additional Information	

# 音声自然降下の知覚的補正に関する予備的研究\*

松井理直

---

## A Study of Perceptual Revision of Declination

Michinao Matsui

### Abstract

This study consists of three experiments, which were carried out to investigate how fundamental frequency declination is perceived by speakers of Tokyo dialect. The results are the following;

- The perception of pitch accents always obeys the categorical perception strictly.
- The perceptual boundary of pitch accent falls down to lower frequency by declination. This result suggests the function of perceptual revision of declination.
- The range of this downslide of the perceptual boundary does not always equals the magnitude of declination. This result means that the effect of declination do not be normalized on the “primitive process” as the general auditory processing.

---

\*本研究は、平成 11・12 年度文部省・日本学術振興会科学研究費 (奨励研究 (A), 課題番号: 11710291) 「日本語動詞形態素の獲得に関する理論的・実証的研究」の援助を受けている。  
*Theoretical and Applied Linguistics at Kobe Shoin* 4, 77–84, 2001.  
© Kobe Shoin Institute for Linguistic Sciences.

本研究は、東京方言に見られる音声自然降下がどのように補正されているかを実験的に検討したものである。その結果、(1) どのような自然降下においても、アクセントの知覚様式は範疇的知覚様式になること、(2) 自然降下によって知覚境界の引き下げが起こること、(3) ただしそのシフト幅は自然降下の割合を十分に補正するほど強くはないこと、が確認された。

## 1. はじめに

近年、東京方言の音調生成モデルでは、より自然な基本周波数の変化を実現するため、音声自然降下(時間と共に  $F_0$  が下降する現象)が考慮されるようになっている。藤崎(1984)によるモデル化、および Pierrehumbert & Beckman (1988)の研究は、こうした定量的なイントネーション・モデルの代表例といえよう。藤崎モデルでは、アクセント成分とフレーズ成分の合成によって  $F_0$  曲線が得られ、音声の自然降下自体はフレーズ成分によってもたらされる。一方、Pierrehumbert & Beckman のモデル(以下 PB モデル)では、自然降下をもたらす要因が複数存在する。全体的な傾斜を作り出す *declination* や、*H tone* と *L tone* を補間することにより得られる下降、アクセント核によって引き起こされる *downstep* などがその代表である。

この音声自然降下という現象は、音調生成にとって重要なものであるばかりでなく、音声の知覚にとってもひとつの興味深い問題をもたらす。Ogura, Wang & Wooters (1989) は、東京方言話者を対象に、「雨／飴」の弁別実験を行った。刺激は単語を単独に提示し、「あ」に対して「め」の周波数を様々に変化させて、「雨／飴」のどちらに理解されるかを被験者に判断させるという形で行われた。その結果、東京方言の高さアクセントの知覚が範疇的知覚に従い、第1モーラの「あ」に対し、第2モーラの「め」が低ければ「雨」に、逆に高ければ「飴」に知覚され、どちらにも理解可能な曖昧な領域は極めて狭いことが示された。Ogura らの研究結果をまとめると、アクセント知覚が範疇的に行われ、その知覚境界(アクセント核直後のモーラの「高音／低音」の基準)はアクセント核を持つモーラの周波数に一致するということになる。では、自然発話のように自然降下を伴うような音声の場合、アクセントの知覚境界はどうなるのだろうか。単語単独の場合と同じく、アクセント核を持つモーラの周波数に一致するのだろうか。あるいは、音声自然降下の影響を受け、知覚境界はより低いほうに移動するのだろうか。こ

の疑問に対する答えは、アクセント知覚の処理過程を考察する上で、ひとつの手がかりを与えるように思われる。前者の場合、アクセント知覚は厳密に局所的に行われていると言えるだろう。一方、後者の場合は、より大局的な処理が関っており、アクセント知覚にとって背景効果である音声自然降下を補正するような処理過程が示唆される。

本研究では、以上の問題意識から、音声自然降下がアクセントの知覚境界にどのような影響を及ぼすのかを確認したものである。なお、本研究は予備的なものであり、詳細な議論は稿を改めて行う予定である。

## 2. 実験 1

### 2.1 実験方法

東京都出身の男性 11 名、女性 9 名 (19 ~ 27 歳) を被験者に実験を行った。実験方法は以下の通りである。まず、音声刺激には「大粒のあめ (雨／飴)」という文章を用いた。刺激音の条件は、「大粒の」の部分に関して、次の 4 種を定めた。

- (1) a. 条件 A: initial lowering (IL) はあるが (おおつぶの)、自然降下のない音声。
- b. 条件 B: IL があり、かつ自然降下を伴う音声。
- c. 条件 C: IL がなく (おおつぶの)、かつ自然降下のない音
- d. 条件 D: IL はないが、自然降下は付随する音声。

全ての刺激音声で、「大粒の」における「の」の周波数を 154 Hz に固定し、「あめ」の部分は 120.9~170.2 Hz まで、1/14 octave ずつ 7 段階で、互いに独立に変化させた。自然降下は、拍内変動を防ぐために、モーラ毎に段階的に変化させた。<sup>2</sup> 1 モーラごとの自然降下の割合は「あめ」の部分における音高変化の割合と同じく 1/14 octave とした。また、各モーラにおける定常部の持続時間は 200msec に保った。音声の大きさについては、調整法により、各モーラにおける音声の loudness が 75dB の pink noise とほぼ同一になるよう intensity を操作した。

<sup>2</sup> この刺激では、音声の自然性が極端に失われてしまうが、1 つの聴覚イベント内で周波数変動が起こった場合、ピッチ知覚そのものが不安定になる可能性があるため、本実験では拍内変動が生じないようにしている。後述の実験 2 では、連続的な周波数変化をつけている

実験手順は以下のとおりである。刺激音はヘッドフォンから両耳にモノラルで呈示した。被験者には、同定実験および ABX 法を用いた弁別実験により「雨」の意味か「飴」の意味かを判断させた。各刺激音声は最終的に 10 回ずつランダムに呈示した。

## 2.2 実験結果

音声自然降下および initial lowering の有無に関わらず、全ての条件で、アクセント知覚の知覚境界は極めて狭いものであった。これは、Ogura らの実験の結果と同一であり、日本語の高さアクセントが範疇的知覚に従うことを確認したものである。

次に、知覚境界の存在範囲について見てみると、自然降下の起こらない条件においては、全ての場合で、「あめ」の「あ」の音高に対し、「め」の音高がほぼ等しくなるところに知覚境界が存在した。被験者の平均を取ると、「飴(平板型)」に知覚される境界は、「あ」の周波数に対し、「め」が 0.25/14 oct 高いところとなった。これに対し、自然降下の起こっている条件では、基本的に、「あ」の音高よりも「め」の音高が低くなるところに知覚境界が存在した。平均で見ると、知覚境界は、「あ」の周波数に対し「め」が 0.21/14 oct 低いところとなった。いずれの場合も、initial lowering の効果は、ほとんど認められなかった。

実験 1 の結果をまとめると、以下ようになる。

(2) a. 自然下降によって、知覚境界は低い周波数域にシフトする

b. 「雨／飴」の弁別が、あくまでも第 1 モーラと第 2 モーラとの関係が重要であり、先行文脈である「大粒の」における低音(第 1 モーラ)の周波数は、「あめ」の低音である「め」の音高関係に全く影響しない。

これらのことから、アクセントパターンを知覚する際に、バックグラウンドである音声自然降下の影響は、なんらかの形で補正されている可能性があるといつてよいだろう。次の問題は、この補正の効果がなにもものかという点である。実験 1 の結果を見ると、音声自然降下の影響によりアクセントの知覚境界がシフトするとはいえ、そのシフト幅は 0.46/14 oct 程度であり、刺激音に伴う自然降下の傾き(1/14 oct)よりは、はるかに小さい値に留まっている。もし、補正の原因が、一般的な聴覚処理機能である streaming (Bregman 1990) のようなものによって引

き起こされるのであれば、知覚境界のシフト幅は更に大きなものになると予測される。

そこで、アクセント知覚境界のシフト幅と音声自然降下の関係を調べるために、以下の実験を行った。

### 3. 実験 2

#### 3.1 実験方法

刺激文は実験 1 と同じく「大粒のあめ」を用いた。刺激音の作成手順は以下の通りである。まず、音声自然降下は段階的なものではなく、連続的に変化するものを用いた。これは、刺激をより自然発話に近づけるためである。自然降下の変化幅は、隣接するモーラにおける母音の中心部の間で一定になるよう操作し、0 oct (自然降下なし)、1/35 oct, 1/28 oct, 1/14 oct の 4 条件を設定した。自然降下そのものは、第 2 モーラの母音開始部以降で実現されるように設定した。各母音の持続時間は先行する子音により若干の変動をつけ、母音単独のモーラでは 160msec で一定にし、最も短い「つ」の [u] の部分では 145msec とした。全ての刺激音声で、「あめ」における「あ」の中心周波数が 154 Hz になるように、「大粒の」の部分における周波数を調整した。また、「あめ」の「め」の部分における中心周波数は、129.5~183.1 Hz まで、1/28 octave ずつ変化させた。以上の条件を満たす刺激音を、Klatt 型の formant synthesizer によって作成した。

実験手順は以下のとおりである。刺激音はヘッドフォンから両耳にモノラルで呈示した。被験者には、PEST 法を用いた弁別実験により、「雨」の意味か「飴」の意味かを判断させた。各刺激音声は最終的に 10 回ずつランダムに呈示した。被験者は、男性 8 名、女性 2 名の計 10 名で、いずれも東京方言話者である。

#### 3.2 実験結果

音声自然降下を連続的にした刺激においても、アクセント知覚の知覚境界は極めて狭い範囲で起こり、範疇的知覚様式が確認された。判断確率 25% から 75% にかけての境界幅は、段階的な自然降下と比較して有意な差は認められなかった。

次に、知覚境界のシフト幅について見てみる。音声自然降下のない場合と、1/35 oct, 1/28 oct, 1/14 oct の各条件におけるシフト幅を計算すると以下ようになる。なお、シフト幅は、「飴」の 75% 知覚判断域を基準に算出したものである。

- (3) a. 自然降下 1/35 oct におけるシフト幅：0.61/28 oct
- b. 自然降下 1/28 oct におけるシフト幅：0.87/28 oct
- c. 自然降下 1/14 oct におけるシフト幅：0.96/28 oct

この結果を見る限り、確かに自然降下の割合が大きくなるにしたがって、知覚境界のシフト幅も大きくなってはいるが、補正の効果には限度があり、ある一定の割合以上にシフトすることはないように思われる。音声自然降下を代償するような知覚上の補正効果は確かに存在するが、その効果は時にかなりの傾斜を持つ音声自然降下の傾斜を常に補正できるほど強いものではないと言えるだろう。

## 4. 実験 3

PB モデルの興味深い点の一つに、H tone と L tone の間に存在するモーラ数によって、その傾斜の割合が異なるという主張がある。彼らのモデルでは、こうした傾斜は H tone と L tone を補間して得られるものであるため、両音間に存在するモーラ数が少ないほど、下降の傾斜の程度が強くなるのである。実験 2 で、補正効果の飽和が見られたのは、この補間によって得られる傾斜の割合の限度を越えてしまったからとも考えられる。本実験では、この効果の知覚に及ぼす影響を簡単に調査する。

### 4.1 実験方法

刺激文は、「ここのあめ」、「大粒のあめ」、「実物大のあめ」の 3 種類を用いた。第 2 モーラにおける母音開始部の周波数は 165Hz に固定し、「あめ」における「あ」の中心部の周波数は各々 155Hz, 153 Hz, 152.1 Hz に設定した。また、第 2 モーラから「あ」にかけての自然降下は、実数値上で (対数的にではなく) 線形的に補間した。PB モデルに従うなら、これにより平板型である「飴」における「め」の理論上の最適値は、全て 150 Hz になることが期待される (知覚上はこれ以上の値であっても「飴」に知覚されてもよい)。最終モーラである「め」の周波数は、140Hz から 160Hz まで 1 Hz ずつ変化させた。各モーラの持続時間は、145msec から 160 msec の範囲に収まるよう調整した。

実験手順は実験 2 と同じく、刺激音をヘッドフォンから両耳にモノラルで呈示し、被験者には、PEST 法を用いた弁別実験により、「雨」の意味か「飴」の意味

かを判断させた。各刺激音声は最終的に 10 回ずつランダムに呈示した。被験者は、男性 8 名、女性 2 名の計 10 名で、いずれも東京方言話者である。

#### 4.2 実験結果

各条件「ここのあめ」、「大粒のあめ」、「実物大のあめ」における「飴」の 75% 知覚判断域は、各々 154.5Hz, 153.3Hz, 152.5Hz であった。PB モデルに従うなら、「飴」の場合には、H tone と L tone の単純な直線補間が行われ、いずれの場合でも 150Hz 付近が判断境界になると予測されるが、それよりは高い数値が得られている。また、各刺激における判断境界も、「ここの飴」と他の条件間との間には有意な差が存在する。

ただし、この実験結果は、必ずしも PB モデルを否定するものではないことに注意されたい。前述したように、PB モデルにおいては、自然降下をもたらす要因が複数存在する。本実験においては、そのうち、H tone と L tone の直線補間という要因のみを考慮したに過ぎない。少なくとも、自然降下の知覚的補正は、音韻表象における H tone と L tone の単純補間から算出されるようなものではないのであろう。

### 5. 総合論議

本研究の実験から、音声の自然降下を知覚的に補正することが可能であること、またその補正はアクセント知覚の範疇的知覚様式そのものには影響を及ぼさないこと、補正には限度があり、時に自然発話に見られるような大きな自然降下を補正するには弱い効果しか持たないことを確認した。

この補正効果の原因は、明確ではない。少なくとも、stream のような一般的聴覚現象のみで説明できるものではないが、知識レベルの音韻表象から精密に計算されるものでもない可能性が高い。ただし、音韻表象からの計算に関しては、限定された要因しか考慮していない。PB モデルにおけるいくつかの要因に関しては、大阪方言における基本周波数の「上昇」とアクセント知覚の関係を調べることにより、その影響を検討することが可能であると思われる。今後、これらの点を中心にさらに検討を行う予定である。



## 参考文献

Bregman, A. (1991). *Auditory Scene Analysis*. The MIT Press.

藤崎博也 (1989). 「日本語の音調の分析とモデル化」. 杉藤美代子 (編) 『講座 日本語と日本語教育 第2巻 日本語の音声・音韻 (上)』. 明治書院

Ogura, M., Wang, W.S.-Y., & Wooters, C. (1989). W.S.-Y.Wang (編) 『発話の基礎』. p.84-85. 研究社

Pierrehumbert, J.B. & Beckman, M.E. (1988). *Japanese Tone Structure*. The MIT Press

杉藤美代子 (1981). 『日本語アクセントの研究』. 三省堂

**Author's E-mail Address:** matsui@sils.shoin.ac.jp

**Author's web site:** <http://sils.shoin.ac.jp/~matsui/>