

スペシャル・セッション [プロソディー生成・知覚・処理] (招待講演) 2—5—12 イントネーションの言語学的理論とその問題点*

○前川喜久雄 (国立国語研究所) kikuo@tansei.cc.u-tokyo.ac.jp

1 目的

言語学的なイントネーション理論の代表としてPB理論 (Pierrehumbert & Beckman 1988) を紹介せよという依頼である。2節で同理論を紹介し、3節では現在のイントネーション研究の問題点につき私見をのべる。

2 PB理論の梗概

この理論は英語を素材としたPierrehumbertの博士論文を萌芽とし[1], のちに日本語(東京方言)への応用がこころみられた[2,3,4]。その基本的な特徴を簡単に紹介する。(文献[5]にもPB理論の解説がある。併読されたい。)

2.1 トーンへの分解

音韻論は離散的言語単位についての記号操作の論であるから、イントネーションの音韻論が達着する基本問題のひとつは、イントネーション(F_0 の時間変化パターン)を、いかにして記号操作に還元するかという問題である。PB理論では発話の要所要所に散在する tone (ピッチに関する抽象的な音韻表示)の連鎖が音声実現規則による補間を受けてイントネーションが生じるとみなす。基本的には autosegmental phonology に立脚した分析であるが、従来、生成音韻論による日本語韻律の分析が表層においてはすべてのモーラ(ないし音節)が tone を担うとみなしていたのに対し、PB理論はより underspecify された表層音韻表示を採用している。

日本語の tone としては語彙アクセント(HL)のほか、韻律構造(後述)の各種境界を表示する機能をもつ boundary tone(L%, H%), およびアクセント句の冒頭ちかくに挿入され、無核アクセント句ではその頂点をなす phrasal H が想定されている。一方、英語の tone には次の3種のカテゴリがある。(a) ストレスをもつ音節に生じて言語情

報の焦点(focus)の所在を示す pitch accent ($H^*, L^*, L^*+H, L+H^*, H^*+L, H+L^*$; この他に文献[1]では H^*+H も。*はその tone がストレスをもつ音節に結合されることを示す), (b) イントネーション句の境界に、冒頭では随意的に末尾では義務的に生じる boundary tone(L%, H%), (c) イントネーション句末尾の boundary tone とその句内部の最終 pitch accent の間にだけ生じる phrase accent (L, H)。

英語には語彙の対立に関与する tone は存在しないが、その代わりに tone連鎖が一種の形態素として機能し、主としてパラ言語的な意味(疑問、不信心、驚き...)を表わす。もちろん tone の付与にさきだってストレスの配置と相対強度が決定されていないなければならない。

2.2 韻律構造

英語において可能な tone の連鎖は文献[1]では有限状態文法として記述されていたが、文献[2,3]では階層的な韻律構造の存在が強調されるようになった。韻律構造は言語ごとに定まる有限の階層数が規定され、接点の再帰的出現は許容されない(strict layer hypothesis)。日本語の場合、階層の接点として、上から順に v (utterance発話), ι (intermediate phrase), α (accentual phraseアクセント句), ω (word韻律語), σ (syllable), μ (mora) が認定されている。 v は発話の冒頭と末尾に boundary tone を導入するとともに、declination(後述のダウンステップとはちがって音韻構造に左右されないゆるやかなピッチ下降)の領域である。 α 接点は支配領域の左右境界に phrasal H と boundary L% を導入する。 α 接点によって支配されるアクセント句はアクセント付与の領域であり、句内部に有核の韻律語が2個以上存在する場合は、もつとも左側のものを除いて他のアクセントは削除さ

* Linguistic study of intonation and its current problems, Kikuo Maekawa (National Language Research Institute)

れる(dephrasing[10]). ι 接点はダウンステップの適用領域を規定する(後述). なお, 日本語には ι 接点によって導入される tone は存在しないが, 英語では前述の phrase accent がこの接点によって導入される[2].

2.3 ダウンステップ (Catathesis)

イントネーションを tone の連鎖とみる観点の萌芽は, はやくアメリカ構造主義言語学のピッチ音素にみとめることができるが, その問題点として指摘されたのが, ピッチにいくつのレベルが必要かという問題である. 構造主義音韻論では通常4種程度のピッチ音素を認定していたが, 単純な段階観では現実の基本周波数の様相を十分に説明することはできず, また音韻論上の一般化をおこなう上でも問題があった. 文献[1]の重要な貢献は, ピッチ・レンジの概念とダウンステップ規則を導入して, 音韻的な tone を H と L の2種に限定しても(むしろ限定すればこそ)広範な F_0 資料を説明可能であることを実証したことであった. 英語では H^*+L 等の two tone accent が契機となってダウンステップが生じ, 後続する tone の実現値を低下させる.

当初英語に関して提案されたダウンステップのアイデアは, Poser によって日本語の分析に応用され[4], P B 理論へと継承された(術語としてはダウンステップのかわりに catathesis が用いられている). 日本語でダウンステップの契機となるのは語彙アクセント(HL)である. 韻律構造中の ι 接点はダウンステップの作用域を指定するもので, ι 領域内に複数のアクセントが存在する場合, ピッチ・レンジは指数関数的に狭まってゆくものとみなしている. 一方, ι 接点の境界ではピッチ・レンジがリセットされる. たとえば発話の一部が文脈上対比のフォーカスをうける場合には, フォーカスをうける要素の直前に ι 境界が挿入される.

ところで P B 理論ではピッチ・レンジは音韻論的な契機によって生じる段階的な変化であるダウンステップ以外にも様々な語用論的要因によって連続的に変化するとみなされている. ピッチ・レ

ンジに連続的な変化が生じうることは, たとえば同一の発話を高低さまざまな声域で発話しても音韻論的な同一性が損なわれないことから明らかとってよいのだが, その適用を適切に制限しないと音韻論的には様々な問題をひきおこす概念でもある(3.2参照).

2.4 音声実現規則

Tone と階層表現からなる韻律構造を物理量に変換する規則群が音声実現規則である. 持続時間についての規則等も必要であるが, 文献[4]では専ら F_0 領域について検討がおこなわれている. 具体的にはピッチ・レンジの下限と上限に対応する抽象的な尺度上で tone の目標値を正規化して表現し, その尺度と F_0 値の対応関係を規則が操作するという方法が採用されている. 計算された目標値は直線補間され, F_0 パターンが生成される.

音声実現規則はその出力が数値である点とならんで「左から右へ」適用される点が音韻規則とちがっている. つまり, 任意の時点で音声実現規則が参照しうる情報は韻律構造のうち当該時点までに処理をうけた部分に限定される.

3 イントネーション研究の問題点

3.1 局所性

P B が発話全体のイントネーションを tone という形で局所的に指定される音韻特徴に帰納するのに対して, たとえば藤崎モデルはフレーズ成分という非局所的な音韻特徴を想定する. アクセント成分も P B の観点からすれば非局所的である. 一般に言語学には局所性を指向する傾向があるが, 単に好みの問題ばかりではなく, 言語事実のなかに音韻特徴の局所性を要求するものが存在していることは事実である. 例として日本語の熊本方言をとりあげよう. この方言は語彙アクセントをもたない無アクセント方言であるが, 東京方言などのアクセント句に相当する韻律単位ははっきりと存在している. 熊本方言のアクセント句(仮にこう呼ぶ)のピッチ形状は, 概略両端がひくく内部に1箇所だけピークが存在するというものであるが, そのピークのタイミングについては厳密な規定がないとおもわれる. FIG.1 に模式化して示し

た疑問詞疑問文「ナンノミュ？」（＝何が見える？）の多様なイントネーションはすべて現実に観察されるものであり知覚実験によっても適格性が実証される[6]。ここで興味ぶかいのは、これらのイントネーションに共通する音韻論的な特性をどう捉えるかという問題である。LHLHという4個の tone がこの順番で並び、そのうち3個は実現位置が規定されているが、句のピークをなすHについては句内部であればどこで実現してもよいわけである（A6は2番目のLが削除されて生成される）。局所的な理論では実現位置の不確定な tone の存在を仮定しても問題は生じない。一方、非局所的なモデルでは、これらのイントネーションを生成すること自体は可能であろうが、その場合たとえば1と5とは別種のフレーズ成分を想定せざるをえず、共通性の把握という目的には難がある。この例で局所性の得失に関する議論がかたづくわけではないが、局所的な理論の長所が端的に発揮されている例といえるだろう。

3.2 韻律構造の問題点

韻律構造と統語構造との写像関係についてはPBはほとんど何ものべていない。一般に、この領域はイントネーション研究のなかでも研究がされている領域である。たとえば日本語に関してしばしば言及される左/右枝分かれ境界におけるピッチのたてなおしに関しても、その基本的な性質は言語学の観点からはまだ十分に解明されているとはいえない。もっとも単純な分析は統語構造におけるこの種の境界が韻律構造の i 境界に写像され、したがってピッチ・レンジのリセットが生じるとするものであるが[7]、実際に測定をおこなってみると、左/右枝分かれ境界によるたてなおしの効果と併存して、当該境界をまたぐふたつのアクセント句のピーク間にダウンステップの効果が認められる発話（つまり無核アクセント句＋有核アクセント句にくらべて有核＋有核の場合に第2句のピークが統計的に有意に低い発話）をおこなう話者がすくなくない。結果として、計測されるダウンステップの強さ（ピッチ・レンジの狭帯係数）には統語構造と相関するかなり大きな

変動がみとめられることになる。同様の相関は、東京方言の疑問詞疑問文と単純疑問文との間にも明白な例をみることができる[8]。（英語の例について文献[9]参照）

ところで、統語構造と相関するイントネーション上の差異をすべて韻律構造に反映させようとする、上述の strict layer hypothesis は、たとえば一部の接点の再帰出現を許容するという形で、すくなくとも部分的な変更を余儀なくされる。そのため、PB理論の提案者達はこの種の変動を韻律構造に起因するピッチ・レンジの段階的な変動とはみなさずに、語用論的要因によるピッチ・レンジの連続的な変動とみなす立場をとっている。この立場にたつての談話イントネーション研究もおこなわれているが[11]、現在までに報告されたデータを首尾一環して説明可能な語用論的な原則は提示されておらず、現状では語用論が理論上のゴミ箱として利用される危険性がある[12]。（なお、PB理論にもとづいてF₀データから韻律構造を推定する際の問題点について[10]参照。）

3.3 韻律特徴と分節音の調音

PB理論にかぎらず現在のイントネーション研究ではF₀特徴を分節音の特徴と分離したうえで考察することがおおい。これが有効な作業仮説であることは経験的にみて確実であるが、実際の音声生成過程において両者が本当に独立しているかというところではない。FIG.2に示すのは「次郎がお爺さんに会った」および「次郎がお婆さんに会った」の下線部のプロミネンスを3段階（一強調なし(N)、普通に強調(M)、特につよく強調(S)）に変化させて繰り返し発音した男性1名のデータを用いて、下線部にふくまれる長母音/a/, /i/および「お婆さん」中の接頭辞/o/のフォルマント周波数を測定した結果である。プロミネンスの増大に伴って、/a/の調音はより広く、/i/の調音はより狭く変化する。また/i/はプロミネンスが最大になると前よりに、/o/は強調があると後ろよりに変化する。これは各母音に固有の音韻特徴([OPEN],[CLOSE],[FRONT],[BACK])を強調する方向への変化であり、調音変化が喉頭と舌と

の機械的なカブリング(だけ)に起因するものではないことを示唆している。FIG.2に観察される変化は音韻の範疇的变化をひきおこすほどのものではないが、パラ言語的な音声情報の知覚に際しては、この種の調音上の変化が(おそらくは喉頭音源の性質の変化とならんで)何らかの役割をはたしている可能性がある。各種韻律特徴と分節音の関係は今後の重要な研究課題である[13]。

文献

[1] Pierrehumbert, *The phonology and phonetics of English intonation*, MIT diss., 1980.
 [2] Beckman & Pierrehumbert. *Intonational structure in Japanese and English*, *Phonology Yearbook*, 3, 1986.
 [3] Pierrehumbert & Beckman. *Japanese tone structure*, MIT Press, 1988.
 [4] Poser. *The phonetics and phonology of tone and intonation in Japanese*, MIT diss., 1984.
 [5] 藤村「音声・音韻研究の展望」講座日本語と

日本語教育 2, 明治書院, 1989.

[6] Maekawa. *Intonational structure of Kumamoto Japanese: a perceptual validation*, *Proc. ICSLP94*, 1994.
 [7] Selkirk & Tateishi. *Syntax and downstep in Japanese*, in *Interdisciplinary approach to language*, (Georgopoulos & Ishihara, Ed.), Kluwer, 1991.
 [8] Maekawa. *Perception of intonational characteristics of WH and NON-WH questions in Tokyo Japanese*, *Proc. 12th ICPPhS*, Aix-en-Provence, 1991.
 [9] Ladd. *Intonational phrasing: the case for recursive prosodic structure*, *Phonology Yearbook*, 3, 1986.
 [10] Maekawa. *Is there dephrasing of the accentual phrase in Japanese?*, *OSU WPL*, 44, 1994.
 [11] Hirschberg, & Pierrehumbert. *The intonational structure of discourse*, *Proc. 24th Ann. Meeting of the Association for Computational Linguistics*, 1986.
 [12] Ladd. *In defense of a metrical theory of intonational downstep*, in *The phonology of tone*, (van der Hulst & Snider, Ed.), Mouton de Gruyter, 1993.
 [13] Sproat & Fujimura. *Allophonic variation in English /l/ and its implications for phonetic implementation*, *Journal of Phonetics*, 21, 1993.

Figure 1. Schematic representation of six well formed intonation contours of WH question sentence /nanno miyu?/ (=What can be seen?) in Kumamoto Japanese. The phrase medial H can be realized anywhere within the phonological phrase to which it belongs.

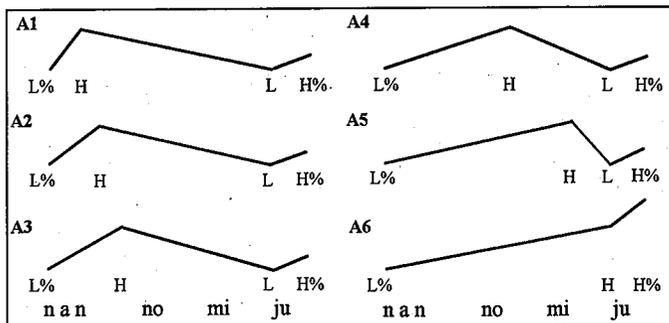


Figure 2. Effects of different degrees of prominence on the formant frequencies of the vowels /i/, /a/ and /o/ in Tokyo Japanese (An adult male subject). The focal condition of each data point is shown by plotting symbols (No focus, Moderate focus, and Strong focus). Also, the fifty percent probability concentration ellipses are used to show the overall distributions of the three data clouds.

