

# 韻律構造境界声調表示の類型論

## Typology of Prosodic Structures: Multiple Tonal Representation

児玉望

Kodama Nozomi

### <プロジェクト研究概要>

日本語諸方言をはじめ、アジアの諸言語には、韻律構造階層（イントネーション句、音韻句、音韻語）の境界をいずれも音節声調によって表示する言語がある。特に、音韻語がこの音節声調配列によって弁別されるタイプの日本語諸方言では、重層的なピッチ変化の合成が観察され、声調変化の制約も強いが、声調による語の弁別のない韻律体系（熊本方言や朝鮮語ソウル方言）では、韻律構造境界の表示とイントネーション弁別のみを声調変化が受け持っており、このような境界表示の体系が観察しやすく、より複雑な弁別的体系の分析にも洞察を提供する。本研究では、フィールドワークによっていくつかの言語の音声コーパス資料を作成し、音節声調分析という共通の方法によって音韻論的解釈を施して韻律構造を解析し、これらの表示がどのような声調によって実現されるかのマークアップ付きで公開資料を作成することを目的とする。対象方言と分担は、朝鮮語ソウル方言（福澤清、博士前期課程 後藤祐司＝平成21年6月まで留学予定）、熊本県中部・北部方言（坂梨瑠璃・児玉望）が分担し、浅田和泉は、日本語韻律構造を日本語学習者がどの程度知覚するか、日本語教育において、このような境界表示データを用いた音声資料がどのように活用できるかの研究を担当する。

### 1. 韻律構造階層

韻律構造 the prosodic structure は、生成音韻論において Selkirk (1984) に集約された構文論と音韻論の関係の全体的な修正を伴う音韻表示理論で仮定された階層構造の一つで、最小単位を音節とし、厳密階層仮説 Strict Layer Hypothesis と呼ばれる、上位の構造境界は常に下位の構造境界でもある（が逆は真ではない）という制約に従うような、超分節的構造、つまり、ある長さをもつ分節音のまとまりの諸段階をいう。言語の音韻表示は、単なる分節音のランダムな連続として機能しているのではなく、「語」なり「句」なりの構造を形成している。これらの「語」や「句」は、意味と密接に関係するが、たとえば、「語」という階層自体が直接に何かの意味に結びつけられるわけではない。名詞も副詞も動詞もみ

な「語」である。「意味のまとまり」としての「語」は、構文論の単位でありその構文論上の役割によって品詞にカテゴライズされるが、音節連続として品詞に関わらず共通の性質（たとえば語中、語末といった音韻環境による異音）をもつ「語」という単位は、音韻論上の単位として（構文論の語とは区別して）定義しうる、という考え方である。構文論上の「語」（たとえば複雑な整数を表す数詞）は、必ずしも音韻論上の単一の語を構成しないし、音韻論上の「語」が構文論上も「語」として機能しうるわけではない。一方、疑問のイントネーションは「文」の発話でも「語」の発話でも共通に現れることができ、この場合、音韻論上は同一の階層を構成していると考えねばならない。このような、独自の構造をもつものとして音韻構造を仮定すると、生成文法理論での表層構造と音韻表示との間の写像関係は、初期の理論で考えられていたものよりかなり複雑になり、同一の表層構造の音声的実現に、音韻構造の異なる複数の形式があり、場合によっては異なる意味解釈を受ける、ということになる。

Selkirk (1984)では、音韻構造の要素として韻律構造のほかに強弱のリズム構造が仮定されており、強弱アクセントの分布や強勢削除によって「構文論上の語や句」が表示される英語のような言語では、韻律構造としては音節とイントネーション句（ピッチ変化で定義されるまとまり）を仮定するだけでよいとし、韻律構造の役割は相対的に小さくなっている。これに対して、リズム構造も含めて（あるいは特定のリズム構造に結び付けて）音韻論上の語や句を定義する、という一般化もあり、Nespor & Vogel (1986) はこのような一元的な韻律構造階層を提案している。

韻律構造でリズムを統一的に扱にくい一つの理由は、厳密階層仮説が要求する各構造が、音節境界に一致した両端をもつ（境界表示された）まとまりでなければならないのに対し、強勢音節が境界表示機能をもつと考えるためには何らかの位置の制約を仮定せねばならず、語頭なり句末なりの位置固定型の強勢でない限り、境界が必ずしも一意的に決定できないからではないかと考えられる。これは特に、2音節から成る foot のような、しばしば（構文論上の）語よりも小さな、従って意味を手がかりにはできないような単位を仮定しなければならない階層において問題を孕んでいる。また、「(量的) 強勢」に限らず、たとえば段階声調によるピッチアクセントであっても何らかの卓立した音節(またはモーラ)を選ぶような構造では同じ問題が生じうる。Selkirk(1984)で「フォーカス句」を定義するのはこのようなピッチアクセントであり、「境界」に関しては統語構造のそれに依存せざるをえない。

これに対し、児玉(2008)では、鹿児島方言など日本語の多くの方言において、さまざま

な階層が、音節の部分から音節連続までのさまざまな曲線声調（ピッチ変動パターン）で表示される韻律構造を持っていると主張した。鹿児島方言の場合、音韻論上の語が末音節周辺でのピッチ下降パターン、音韻論上の句が冒頭語でのピッチ回復パターン、イントネーション句が最後の音節での音節（モーラ）声調パターンによってそれぞれに境界付けられている。音声の時間的なまとまりを表示するためにはピッチの時間的連続と不連続は効率的であると考えられる。一方、発話中のすべての音節が3つの構造のいずれかによって指定された曲線声調を実現しなければならないため、たとえば英語のイントネーションによるピッチ変動と比べると、その変異の自由度はかなり低いと考えられる。一方、東アジアや東南アジアに数多く見られるいわゆる音節声調言語の場合には、弁別的な声調が各音節のピッチ変動を指定してしまうため、音韻論上の語や句の境界表示についてはおそらく曲線声調以外の韻律的特徴（リズムや段階声調）が関与していると考えねばならない。

しかし、どんな手段であるにせよ、音声言語が普遍的に、知覚可能な特徴によって境界づけられたいくつかの階層的構造をもっていることは、暗黙に了解されると考えていいと思われる。初めて聞く外国語や方言の聞き取りで直面する困難は、おそらくは、音節をはじめとするこれらの構造の境界を見つけられない、ということであろう。これらの韻律構造階層をいくつ想定すべきかは言語ごとに異なる可能性がある。また、それらの韻律構造が（文法構造の異なる）言語間に対応していると考えられるかどうかは、たとえば分節音を用いる場合（日本語の終助詞など）も含め、言語構造全体の問題になるだろう。このような領域は、各個別言語についても先行研究が乏しい分野であった。一つには、伝統的に、「語」以上の単位を構文論、「語」未満の分野を音韻論の守備範囲であると見なす記述スタイルが支配的であったこと、また、実用語学においては、語以上のレベルに関しては音声言語よりは文字言語の習熟が優先されたこともあろう。ただし、近年のデジタル音声技術の日常化により、危機言語をはじめとする無文字言語を含めた言語資料として、音声記録の蓄積や検索可能な形での配布が急速に進行しており、語学教育においても大量の音声教材の日常的な利用が可能になった結果、これらの効果的な利用が求められている。

本プロジェクトは、鹿児島方言タイプの、主として曲線音調による音韻構造階層境界表示を音韻構造類型の一つと見なし、このタイプの言語で各階層がどのように境界表示され、どのような機能を担っているかを、談話音声コーパスに基づいて分析し、これらの境界を分節音表示に加えてどのようにマークアップするかを考える、というものである。対象として、熊本方言のような、ピッチによる（音韻語レベルでの）弁別がなく、ピッチ変動が語あるいは句の境界表示やイントネーション表示のみに用いられており、強勢の機能が限

定的な言語を選び、イントネーション句や音韻句については、それらの境界声調がどんな離散的な体系をなしているかを分析する。つまり、「ラング」の分析であり、分節音の音素の弁別を分析する場合と比べればわかるように、一回性の「パロール」であるコーパス資料のみに基づいて分析することは厳密にはできないのであるが、対象が比較的容易に（機械による）客観的分析が可能なピッチ変動であることと、すべての境界特徴は知覚可能であり表層での音韻実現で観察可能であることから、ある程度の長さのあるコーパスがあれば、少なくともコーパスに現れている実現形については分析案は提示できるのではないかという楽観的な見通しで取り組む。正確な分析のためには言語資料提供者による面接調査が必要であることは言うまでもないが、話者の高齢化が進んでいる場合のように、この選択肢が限られている言語が多いことも考慮する。

## 2. 「無アクセント」言語

本プロジェクトで記述の対象とする、アクセントによる「語」の弁別がないとされる言語の韻律構造について、ここで簡単に整理しておく。このような言語で、韻律特徴は語や句の境界表示機能を担うと考えられるが、それに特化するのではなく、「語」以外のレベルでの韻律特徴による弁別（たとえば句末イントネーションによる弁別）は、もっていてもおかしくない。境界表示機能の実現のためには、境界付けられた構造がある程度固定したパターンを持っていなければならないし、弁別的な可変部と、可変でない部分はある程度明確に区切られているはずである。「ある程度」と留保をつけたのは、固定したパターンとはいっても自由変異的に（必ずしも意味をもたずに）交替する複数の形をもつことは予想されるし、また、一方では下位の音節構造や分節音側から、他方では上位階層の音韻句なりイントネーション句なりからの条件付けによる必ずしも分析的ではない交替が起こりうることも考慮しなければならないからである。ただし、この点は「語」の弁別がある場合でも同様に生じうることであり、複合の組み合わせが少ない分、弁別のない言語のほうが相対的に分析しやすい（境界を見定めやすい）と考える。

もう一点、「語」の弁別がない言語の特徴は、音韻的語としてのアクセント句の境界表示を削除するためのコストが、ある場合に比べて低い、という点である。型の弁別がある場合には、境界削除の結果として生じる構造がどんな型になるかを決定する仕組みがない限り境界削除ができないが、型の弁別がなければ、削除後の連続がやはり音韻的語となる、というデフォルトの条件だけで境界削除が可能である。実際、プロジェクトメンバーの坂梨が修士論文において音声コーパスに基づいて記述した熊本方言における「長距離上昇」

は、このような音韻語境界の削除に音韻句の特徴が組み合わさったものであると考えられる。このような境界削除は、福岡方言のような型の弁別のある言語でも観察されており、必ずしも語の弁別がない言語の特徴とはいえないが、音韻的語の境界表示の有無が機能的対立をなす場合があることを分析にあたって考慮しておく必要がある。

### 3. 分析手順

対象言語は文法構造の概略が既知であり、分節音構成と形態構成とがすでに利用できるという条件下で、形態単位の音韻分析からはじめて上位の句構造を導く、という手順を取る。統語構造が二項枝分かれの統語であり、内心構造(XP→X YP)を中心とした複数のカテゴリーを指定するのに対し、韻律構造階層は各階層をそれ以上カテゴリー分けしないため、構成素分析は、まず最下位の境界を求め、それらの境界の異なる実現のうちどれが上位の境界を兼ねるかを判定する、という、(境界音素を設定するのに似た)線状的な分析となる。

意味の最小単位である形態は、音韻論上は以下の5つのいずれかになると考えられる。

1. 単独で発話を構成できるもの
2. イントネーション句を構成するが、単独で発話を構成しないもの
3. 音韻句を構成するが、単独でイントネーション句を構成しないもの
4. アクセント句を構成するが、単独で音韻句を構成しないもの
5. 単独でアクセント句を構成しないもの

1がいわゆる(単純な)「自立語」であり、2-5は「機能語」や複合語の構成素に相当する。アクセント句を構成するかどうかは、形態の境界で知覚可能なピッチ曲線の不連続があるかどうかで判別できるものと考えられる。境界が判別できない5は、被修飾語や補語(助動詞に対する動詞)など緊密な統語関係にある他の語の形態(同じく5)と共にアクセント句を構成すると分析する。複合語も、成分が共に5であれば1アクセント句を構成すると考えられるが、複合語の中には音韻論上の1語(1アクセント句)とはみなせないものもあることに注意が必要である。

このようにして導かれたアクセント句の境界は、以下のいずれかである。

- A. イントネーション句、音韻句の境界を兼ねる
- B. 音韻句の境界を兼ねる

C. アクセント句のみの境界である。

BとCの違いは、前後いずれかのアクセント句のピッチパターンの違い（句音調を伴うかどうか）として、あるいは、この前後の分節音間の「連声 sandhi」と総称される同化／異化現象によって知覚可能であると考えられる。つまり、「音韻句内部」をドメインとするような音韻現象（たとえば撥音の先行同化）は、Cでは起きるがBでは起きない、と考えられる。Aは、Bのうち、句頭・句末の特徴的なピッチパターンが出現することと、ポーズを置くことができるなどの特徴によって際立っているものであると認められる。アクセント句、音韻句、イントネーション句のすべてに曲線声調が関与する日本語タイプの言語では、イントネーション句を特徴付けるピッチ変動は、局所的なものにならざるをえないと予想される。句末イントネーションと句音調は、それぞれ弁別可能な有限個の実現形から成る弁別音韻体系を構成し、意味弁別の単位としての「イントネーション」は、特定の句末イントネーションと句音調との組み合わせによって実現すると考える。

Aの境界の前後のイントネーション句は、その統語的構成に関わらず語順の交換が可能である。従って、語順が自由な語は、上記の形態の分類のうち1あるいは2に相当すると考えられる。単独で構文論上の句を構成するものとしては、接続詞や文修飾の副詞がこれに相当する。これに対して、5以外の付属語は3あるいは4となる。3は、プロミネンスを付与することのできる付属語や、Bを挟んでホストとなる句と2句構造を構成してホストにプロミネンスを付与する付加語的な付属語（「ばかり」「だけ」など）である。単独でアクセント句を構成しない5の中には、対比フォーカスを置くことが可能な語（否定助動詞「ん」など）もあり、この場合、5が構成するアクセント句全体が境界Bによって区切られることになる。

以上のような点を考慮し、出現するアクセント句の曲線声調を整理してその解釈を検討する。Jun(2005<sup>1</sup>)などソウル方言に関する先行研究が示すように、アクセント句の音節数やその構成、語頭子音などで分類して分析する必要がある。ただし、Cのアクセント句境界に関しては、対象言語（熊本方言・ソウル方言）において、境界の後ろのアクセント句の音節構成に関わらず、いずれもピッチ下降が観察されて境界が明示されていることがわかる。いずれの方言でも音節間での急激な下降が観察できるが、アクセント句末音節における下降調は、ソウル方言の場合はこのアクセント句がイントネーション句あるいは音韻句の境界(AまたはB)の前にあることを示すようである。熊本方言の場合、Cの後では必ずピッチが低い、その下降が前の音節内部から始まるかどうかは必ずしも弁別的ではないよ

うである。アクセント句冒頭でのピッチ上昇のパターンが、どちらの方言においても音韻句の境界を示すと考えられる。熊本方言の場合は、B に先立って、音韻句末の音節に必ず下降調が観察される。

句音調と句末イントネーションの弁別体系の網羅は、本プロジェクトの守備範囲ではないが、限定された音声コーパスの範囲内でも出現する変異については記述しておく必要がある。句音調に想定される変異形としては、プロミネンス付与に関わるものがある。児玉(2007)において、先行研究で「音調句」(音韻句)と「フォーカス句」が同一視されていることを批判し、音調句への分割はプロミネンス付与(対照フォーカス表示)の必要条件であり、音調句はプロミネンス付与の単位となっているだけであると主張した。すべての発話が一個以上の音韻句から成る以上、これらがすべてフォーカス付与されているはずはないからである。どの音韻句にプロミネンスが付与されているかが知覚できるとすれば、その弁別は、句音調に有標・無標の区別があることによると考えるのが自然である。有標・無標の弁別にどのような音声的特徴が関わっているかを判別するためには実験音声学的な研究が必要になると考えられるが、談話コーパスにおいては、どの音韻句がプロミネンス付与されているかが文脈によって判断できるため、「有標」な句音調のデータを集めてその弁別特徴に関する仮説を立てるためにはきわめて有効であると考えられる。従って、プロミネンス付与されていると考えられる音韻句については、必ずマークアップしておく必要がある。同様なことは、句末イントネーションの弁別体系解釈についても言える。

#### 4. 現状での問題点

現段階での問題点は、1 音節から成る形態の音韻解釈である。音韻句であれ、イントネーション句であれ、その境界声調はしばしば句頭・句末の音節の曲線声調として実現する。この位置に現れやすい1音節形態(特に通常は句末位置に現れる付属語)は、その曲線声調がどの階層に関係付けられるかが判断しづらく、このために、アクセント句の境界の有無が決定できない。「韻律構造」は文法的カテゴリーと無関係にそれぞれの階層に一律に実現する、と考えれば、単音節であるが句頭や句末以外の位置にも現れることのできるアクセント句との比較によって、なるべく一貫した基準でマークアップを行なうことが望ましい。ただし、終助詞類のように、出現環境が特殊なものについては、音韻論的解釈以前にデータの収集が求められる分野であり、まずは直観的な分類によってどのような変異が認められるかを洗い出しておく必要があるだろう。このような事例であることの注記として、観察者による境界表示が確定できないもの、というマークアップも必要となるだろう。

## 5. マークアップ付資料の作成

言語の文字表記の歴史は長いが、分節音表記以上の言語音の構造についての表記は、単語の分かち書きのような特定の階層を除けばほとんど慣行化されていない。それらに相当すると考えられる句読点も、「句読法」として、むしろ文法構造を基準としての規範化が行なわれることが普通である。学術目的であれ実用目的であれ、音声言語資料の検索利用に役立つような音韻構造表記の確立を目指すさまざまな試みの一つとして、音声資料付で公開していくべきであると考えられる。

音声コーパスは、フィールドリサーチによって、内容が公開されることに許諾を得たうえで自然な談話を録音することが望ましいが、韓国 국립국어연구원（国立国語研究院）『ソウル方言朗読体発話集 2/3』のような公開資料も用いることができる。

韻律分析の実例は、「凡例」として示し、ここでそれぞれの韻律構造の音形とその弁別体系につき論じた上で、実際のテキストについては、分節音解釈を反映すると考えられる表音的な表記に、韻律構造境界に関する解釈をなるべく簡略にマークアップして表示する、という方法を考える。以下は朝鮮語ソウル方言の例である。

어[디선가 바[보(RF+) 바[보(RF+) [하는 [소리가 자[꾸만 뒤[따라|오는|것|같았다(F%)  
どこからか ばか ばか という 声が ひっきりなしに ついて来るようだった

アクセント句境界(C)は"|" (パイプ) で、音韻句境界はスペースで、イントネーション句境界は、(RF+)などの境界音調表示によって表記する。アクセント句について型の区分のない言語では、アクセント句については境界表示だけで曲線声調の実現は表記できる。音韻句については、句音調による音節間上昇の実現の位置を表記しているが、ここでは句頭分節音の種類により位置が予測できるので、冗長な情報である。ただし、無標な句音調と有標な句音調（対比フォーカスが加わる場合など）とを表記し分ける位置として記号によるマークアップを加えておくことが必要である。イントネーション句の境界を表示する句末イントネーションについては、曲線音調を音節単位（あるいは音節の部分単位）で種別表記しておく。対象言語はいずれも日本語と語順面（韻律構造面）での類型が共通と考えられる言語であるため、対応の音韻句境界ごとに訳を付す程度の簡略なもので十分であると考えられる。このような訳文表記は、利用者が日本語音声の韻律構造について自覚的に反省し、対象言語のそれと比較することができるものであることが望ましい。

音声資料のテキスト化とその韻律特徴のマークアップには、その目的に応じてさまざま

な方法がある。たとえば、ピッチの変動パターンをみの注記法としての INTSINT のように、より音声表記に近いレベルであって、未知の言語や音韻解釈の行なわれていない言語にも適用できるものと、AM(autosegmental metrical)理論の音韻解釈に依拠した ToBI(声調と分割指標 Tone and Break Indices)のように、音韻構造解釈の適用例としてのものがある(Jun 2005<sup>2</sup> 参照)。本プロジェクト研究で目指しているのは、後者のタイプのものである。日本語と朝鮮語も ToBI による韻律構造表記がすでに行なわれているが、本プロジェクトでは異なる音韻解釈モデルを採用している。たとえば、Jun(2005<sup>1</sup>)で言う朝鮮語のアクセント句は本プロジェクトでの音韻句に相当し、ピッチではなく境界の主観的判断に基づく Break Index の側(tier)に「音韻論上の語」を位置づける二本立ての分析となっている。日本語の場合も同様の分析である。このような音韻解釈を実際のデータと共に公開することが音韻論研究では一般化し始めているが、コーパスデータのマークアップ付きの公開は、その意味で、音韻解釈仮説提示とデータ公開・オープンソース化という今後一般化していくと思われる方向を意識したものである。

## 6. 今後の展望

本プロジェクト研究は、韻律構造のすべてをピッチによって境界表示する、という類型を仮定し、例証するものであり、今後、対象言語を同様な類型にあてはまると考えられる他の言語に拡大していくことを意図している。たとえば、南アジアの多くの言語（いわゆる「インド英語」を含む）が同様の枠組みで分析できるのではないかと考えている。

この類型に対置されるべき他の類型との関係についてもここで簡単に考察しておく。上述のように、韻律構造階層は、厳密階層仮説という、下位の階層が上位の階層の領域に完全にかつ排他的に包含されるような（上位の境界が常に下位の境界でもあるような）階層モデルを採用している。本プロジェクト研究が対象とする類型は、音韻語に相当するアクセント句を曲線音調の配列で定義されるものとし、3つの階層がすべてピッチで定義される構造であると考えるのであるが、発話全体がある長さをもつ連続に加わるピッチのパターンによって三段階に分割される、という点で、この仮説によく当てはまっている。しかし、すべての超分節的構造がこのような階層を成すわけではない。たとえば、韻律構造の最下層に位置づけられる「音節」が下位のその構成素との間にもっている関係は、厳密階層仮説には従わない。たとえば、閉音節言語では CVCVC のような連続が、2つの閉音節（CVC+CVC）と解釈されることがよくあり、真ん中のCは2つの上位構造に帰属することになり、厳密階層仮説を満たさない。同様に、英語の（主）強勢や、「pitch accent」（段

階声調の点でプロミネントな強勢)のような、中核となる要素で定義された構造は、それだけでは境界を導くことができない。ToBIの二本立てや、Selkirkのリズム構造との組み合わせによる音韻理論も、このように、境界表示機能が曖昧な構造階層を解釈するための道具立てであるようにも見える。ToBIは、ピッチのみで境界表示している言語に適用すると、冗長性のある、むしろ音声表記的な性格を帯びるようにも見える。

言語の音韻構造に何らかの普遍的な要素があるとすれば、それはピッチ変動やリズムといった表面的な違いには左右されないもっと抽象的な何かだろうと考えられる。とすれば、開音節が比較的多く、ピッチの変動が聞き取りやすい言語にみられるピッチのみを用いた構造の解釈にはそれに合った方法があつてしかるべきであろう。同じ方法が同様な特徴をもつ複数の言語に適用できたなら、言語普遍の可能性のある要素として、他の類型での対応を求める、というような道筋を考えている。その過程で、聞き取り用の語学教材開発のヒントがあれば幸いである。

#### 参考文献

- Jun, Sun-Ah (2005<sup>1</sup>) “Korean Intonational Phonology and Prosodic Transcription” in Jun ed. (2005<sup>2</sup>)
- Jun, Sun-Ah (2005<sup>2</sup>) ed. *Prosodic Typology, the phonology of intonation and phrasing*. Oxford University Press.
- Nespor, Marina, and Irene Vogel (1986) *Prosodic phonology*. Dordrecht: Foris.
- Selkirk, Elizabeth O. (1984) *Phonology and Syntax: the relation between sound and structure*. The MIT Press.
- 児玉望(2007)「音調句と日本語韻律構造」『熊本大学言語学論集6』1-22.