

概念からのプロソディーを含む発話生成

吉本 啓* 小林 昌博* 佐藤 滋* 風斗 博之†

*東北大学大学院国際文化研究科 †東北学院大学教養学部

1 はじめに

言語処理技術を応用した文の生成や音声処理技術を用いた発話生成は広く研究され実用化も進んでいるが、両者の潜在的な能力を生かして意味等の違いによって発話のプロソディー (韻律) を生成し分ける研究は進んでいるとはいえない。これは音韻情報・統辞情報・意味情報などの異なる部門に属する情報の構造上のミスマッチや、また異なる部門からの情報の統合の問題を解決するための理論的な研究が不十分であったためと考えられる。本研究では日本語を対象として取り上げ、主として Head-driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994) を用い、音韻情報を含む辞書・文法規則にもとづいて、発話に対応する音韻情報を生成するための手法について提案する。この統合的な発話生成モデルは上記の音韻構造と意味・統辞構造の間のミスマッチの問題を解決し、また同時に文が修飾構造の違いによる曖昧性を持つ場合に適切な音調の生成を可能にする。

2 日本語のプロソディー

2.1 アクセント句

日本語のプロソディーの形成において中心的な役割を果たすのはアクセント句である。アクセント句はアクセント核をになうことが出来、その前後にはポーズが挿入され、またその終端部分では音の弱めや無声化が行われる。アクセント句は1個の自立語と、それに後続する0個以上の付属語とから構成される。アクセント句の持つ**アクセント情報**—有アクセントか無アクセントか、また有アクセントの場合はどのモラがアクセント核を有するか—は、右端の付属語の持つ**アクセント関数**に対して、それよりも左側の語列のアクセント情報を適用するという操作を再帰的に行うことで得られることが知られている (句坂・佐藤 1983)。

例文 (1a) に文のアクセント句への分割の例を、

また (1b) に同じ文の意味のまとまりにもとづく分析を示す。

- (1) a. [s [s [NP [NP [ADJ きれいな] [N 家]] [P が]] [V 出来]] [AUXV -ます]]
 b. [AP きれいな] [AP 家 が] [AP 出来 -ます]

このように、一般にアクセント句への分割と直観的な意味のまとまりにもとづく構造 (句構造) とは一致しない。これは、意味にもとづく構造が、自立語・付属語の区別にかかわらず修飾-被修飾の関係 (例えば、「きれいな」と「家」の間の) を優先するのに対して、アクセント句はつねに1個の自立語プラス0個以上の付属語という構成を取るからである。本稿ではアクセント句にもとづく文の分割を**P-構造**、直観的な意味のまとまりにもとづく構造を**C-構造**と呼ぶ。

2.2 メトリカル・ブースト

Kubozono (1987) は、日本語のプロソディーがC-構造の影響を直接受けることを明らかにした。すなわち、統計にもとづく音声実験によって、2つの名詞修飾語を持つ句のうち、

- (2) a. [あおい おおきな メロン]
 b. [[[あおい レモン] の] におい]

例 (2a) のように右枝分れの構造を持つ句の2番目の修飾語 (ここでは「大きな」) が例 (2b) のように左枝分れの構造中の2番目の修飾語 (「レモンの」) よりも基本周波数に関して有意の高さの差を見せることが明らかになった。この現象は**メトリカル・ブースト**と呼ばれる。

3つの名詞修飾語を持つ句の場合も同様にして、右枝分れする部分木の中の修飾語がそうでない語よりも高くなる。しかも、二重の右枝分れの左端が1つの修飾語に重なる

(3) [A₁ [[A₂ A₃] H]]

の A₂ の位置の修飾語が最も高いピッチを持つ。

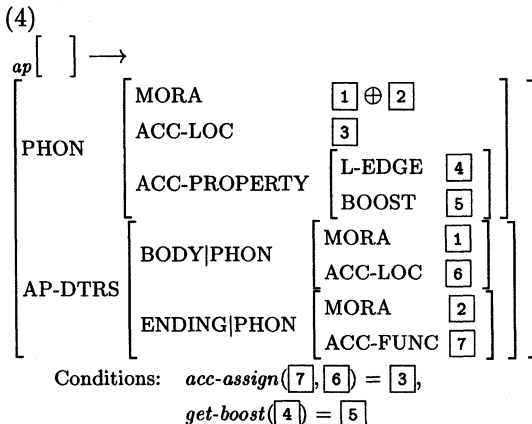
崔・佐藤・関 (1995) は連用修飾句の修飾に関して曖昧な文の理解に際して連用修飾句のピッチの高さが関与していることを明らかにした。従来の音韻論では韻律構造 (P-構造) は概ね自律的なもので、統辞構造 (C-構造) からの影響はきわめて限られたものと考えられてきた。上記の一連の研究は、後者の影響が考えられていたよりもずっと直接的であることを示している。こうして、C-構造からの情報を P-構造に伝えるためのインタフェースの形式化が求められることになる。

3 HPSG にもとづく解決

3.1 基本的な枠組み

前節で指摘した C-構造と P-構造とのミスマッチおよび前者から後者への情報の伝播の問題は、HPSG のタイプ付きユニフィケーション・フォーマリズムを用いて簡潔・効率的に解決出来る。ここで提案する HPSG 文法は標準版に対して、P-構造を構成するための素性 P-STR(UCTURE) を付け加えたものである。

素性 P-STR の値は、アクセント句を表すタイプ *ap* のリストからなる。*ap* は以下の定義を与えられる。



素性パス PHON(OLGY)|MORA の値によって、韻律を除く音韻情報をモラ (拍) のリストの形で表す。あるアクセント句の MORA の値は、そのア

クセント句の右端の付属語 (句末語と呼ぶ) とそれよりも左側の語列 (句幹と呼ぶ) との MORA の値を連結 (⊕ で示す) することによって得られる。素性 ACC(ENT)-LOC(ATION) の値は、自然数の場合は有アクセント句のアクセント核の位置するモラの場所を示し、0 の場合は無アクセントであることを示す。この値は、句末語のアクセント関数についての情報を表す ACC(ENTUAL)-FUNC(TION) の値と句幹の ACC-LOC の値との組み合わせによって再帰的に決定される。

素性 P-STR の値は図 1 の P-構造の原則によって形作られる。これを例文 (1a, b) の主格名詞句「きれいな家が」に適用して得られる P-STR の値を図 2 に示す。これを形作る過程を C-構造 (素性 DAUGHTERS の値) に相当する木によって示したものが図 3 である。図 2 と 3 の間で、素性構造を指示するタグの数字は一致している。

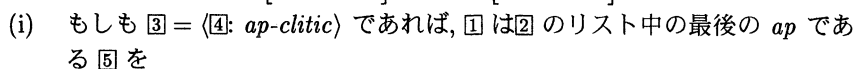
図 3 に見るように、「きれいな」と「家」の P-STR の値をまとめる段階では、これらがともに単独のアクセント句となりうることから、図 1 の P-構造の原則の内の (ii) が適用されて、両者を単に連結したものが「きれいな家」全体の P-STR の値となる。しかし、これにさらに格助詞「が」を付加する段階では、これが付属語であるので (P-STR の値は *ap-clitic*)、(i) が適用されて、句全体の P-STR の値は「きれいな」と「家が」に相当する 2 つの *ap* から構成されるリストとなる。こうして P-STR の値として、C-構造と並行してこれとは食い違う P-構造を構成していくのである。

こうして形成される *ap* はさまざまな音韻規則の適用される領域である。例えば (4) のようにアクセント情報を計算するための *acc(ent)-assign(ment)* はタイプ *ap* を評価する際に誘発される。

3.2 C-構造から P-構造へのインタフェース

2.2 節で解説したメトリカル・ブーストの現象は、C-構造の枝分れについての情報を P-構造に伝えることで解決することができる。このために LOCAL 素性の下に新たな素性 R(IGHT)-BRANCH(ING) を導入し、その値を枝分れの原則によって形作ることとする。

この原則によって、ヘッドとアジャンクトからなる句について、ヘッドが *phrase* である場合、その

$$M: \begin{bmatrix} \text{P-STR} & 1 \end{bmatrix}$$


(ii) それ以外の場合 (すなわち, $\boxed{3} = \text{list}(ap)$ である場合), $\boxed{1} = \boxed{2} \oplus \boxed{3}$

$$pp \left[\begin{array}{l} \text{P-STR} \left\langle \begin{array}{l} \text{①} \left[\begin{array}{l} \text{PHON|MORA} \langle \text{きれいな} \rangle \\ \text{AP-DTRS} \left[\begin{array}{l} \text{BODY} \quad \text{③[PHON|MORA} \langle \text{きれい} \rangle \rangle \\ \text{ENDING} \quad \text{④[PHON|MORA} \langle \text{な} \rangle \rangle \end{array} \right] \end{array} \right] \\ \text{⑥} \left[\begin{array}{l} \text{PHON|MORA} \langle \text{いえが} \rangle \\ \text{AP-DTRS} \left[\begin{array}{l} \text{BODY} \quad \text{②[PHON|MORA} \langle \text{いえ} \rangle \rangle \\ \text{ENDING} \quad \text{⑤[PHON|MORA} \langle \text{が} \rangle \rangle \end{array} \right] \end{array} \right] \end{array} \right\rangle \end{array} \right]$$

Diagram (10) illustrates the hierarchical structure of the sentence in (9). The root node is $pp[p\text{-str} \langle [1], [6]:ap \rangle]$, which branches into $np[p\text{-str} \langle [1], [2] \rangle]$ and $p\text{-wd}[p\text{-str} \langle [5]:ap-cl \rangle]$. The np node branches into $a\text{-wd}[p\text{-str} \langle [1]:ap \rangle]$ and $n\text{-wd}[p\text{-str} \langle [2]:ap \rangle]$. The $a\text{-wd}$ node branches into the AP $きれいだ$ (with $_{AP}$ above it). The $n\text{-wd}$ node branches into the AP $いえ$ (with $_{AP}$ above it). The $p\text{-wd}$ node branches into the particle $が$ (with $_{P}$ above it).

— 399 —

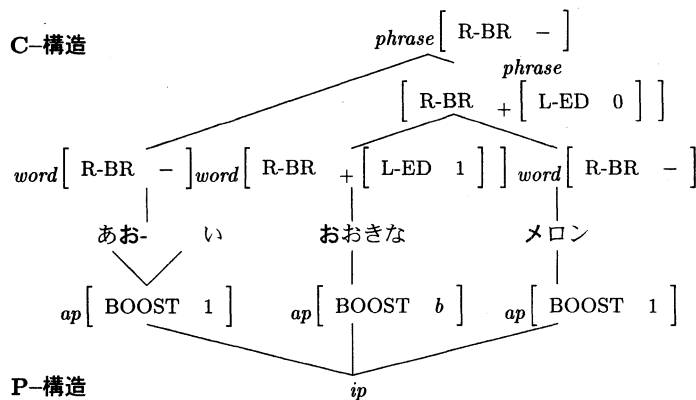


図 4: 右枝分れ構造の BOOST の値

R-BRANCH の値を + とし、それ以外の場合は - とする。+ の場合のみ、値はさらに L(LEFT)-EDGE を素性名とし、値を 0 とする素性構造を持つとする。親の R-BRANCH の値が + である場合は、その L-EDGE の値よりも 1 つ大きい値を子のアジャンクトの L-EDGE の値とする。親の R-BRANCH の値が - の場合は子のアジャンクトのそれも - である。(4) に規定するように、L-EDGE の値にもとづいてアクセント句のメトリカル・プーストの有無およびその程度を表す BOOST の値を決定する。L-EDGE の値が 0 および未規定、1、2 のそれぞれの場合、BOOST の値は 1, b , b' (ただし、 $1 < b < b'$) となる。L-EDGE の値は自立語の定義によって、SYNSEM の下の通常の統辞情報から P-STR の下の ap の中へと伝えられる。

以上の方法によって、例 (2a) の句が分析される過程を図 4 に示す。2 番目の名詞修飾語「大きな」に相当する ap は BOOST の値として b (> 1) を持ち、これによってメトリカル・プーストを受けることが説明される。

これに対して、例 (2b) の句の「レモンの」の BOOST の値は 1 となり、メトリカル・プーストの起きていないことが示される。また、3 つのアジャンクトを持つ句に対してそれぞれ適切な BOOST の値が枝分れの原則によって与えられる。特に (3) の A_2 に対しては、L-EDGE の値が 2 となるので他のメトリカル・プーストを有するアジャンクトの値 b よりも高い b' が与えられ、Kubozono (1987) の実験結果に一致している。

なお、ここで述べた方法について詳しくは Yoshi-

moto (2001) を参照されたい。

4 結び

本発表では HPSG を拡張して、直観的な意味のまとまりと食い違う音韻構造を生成する手法について提案した。最近のプロソディー研究によって、発話のプロソディー生成のためにはこのような離散的な音韻情報で十分であることが知られている。この方法による HPSG の文法規定と音声合成技術を組み合わせ、概念 (論理式) からプロソディーを含む音声までを一貫した手法により生成する実験システムを開発中である。

参考文献

- 崔英淑・佐藤滋・関光準 (1995) 「日本語と韓国語における統語的あいまい性の韻律的解消の分析」『日本音響学会講演論文集』2-P-4.
- Kubozono, H. (1987) *The Organization of Japanese Prosody*. Edinburgh University Ph.D. thesis.
- Pollard and Sag (1994) *Head-Driven Phrase Structure Grammar*. CSLI.
- 句坂芳典・佐藤大和 (1983) 「日本語単語連鎖のアクセント規則」『電子通信学会論文誌』J66-D.
- Yoshimoto, K. (2001) "A Bistratal Approach to the Prosody-Syntax Interface in Japanese." In: R. Cann et al. (eds.), *Grammatical Interfaces in HPSG*. CSLI.