

日本語単語連鎖のアクセント規則

正 員 匂坂 芳典[†] 正 員 佐藤 大和[†]

Accentuation Rules for Japanese Word Concatenation

Yoshinori SAGISAKA[†] and Hirokazu SATO[†], *Regular Members*

あらまし 日本語テキストからの音声合成に必要な韻律情報(アクセント, イントネーション)の生成規則を確立するため, アクセント句を構成する単語の性質からアクセント核の位置を決定する規則を検討した。本規則化では従来の個別的な記述の整理, 分析を行い, 単語間アクセント結合において, 後続単語のアクセント属性として結合アクセント価とアクセント結合様式を提案する。これにより付属語アクセント結合の統一的な規則化を行うと共に, 複合単語についても同様の規則化を図った。また, これらの規則の妥当性を示すため, 自立語に付属語が複数個連なる文節3445文節, 複合単語4877語を用いた推定実験を行った。この結果, 各々98.3%, 95.4%の正解率が得られ, 本規則の有効性が確認された。さらに, これら推定実験結果の分析から, アクセント決定には句を構成する語の性質, 句の構造, 語の用法と意味等が種々に反映されることが示された。

1. ま え が き

文字で表わされた任意の文章を音声で出力する, いわゆるテキストからの音声合成は, 規則による音声合成の一つの最終目標であり, 各国で研究が盛んに行われている^{(1),(2)}。文字系列のみを入力として音声を生成するためには, まず文字系列を適切に解析し, 音韻情報と韻律情報を正しく得ることが必要である。これらの情報は言語的性質とその言語の音声的性質を強く反映しており, 言語によってその様相は大きく異なる。

日本語の場合, 漢字が多く用いられるため, まず, その読みがな(音韻表記)を与えることが必要である。また, 韻律情報として, 単語, 文節, アクセント句(1つ以上の文節からなるアクセント的まとまり), 呼吸段落(一息で発声されるまとまり)といった区分, およびアクセント核の位置などを知る必要がある。ここでは, このような韻律情報生成規則化の第一歩として, アクセント句を構成する単語の性質からアクセント核の位置を決定する問題を取り扱う。

これまで, 日本語アクセントに関しては, 主に言語学の分野で研究が盛んであるが^{(3)~(9)}, 工学的な立場からの研究は少ない^{(10)~(14)}。種々のアクセント辞典^{(7)~(9)}に

も示されているように, アクセント核位置は, そのアクセント句を構成する単語それぞれのアクセント型や単語結合時の諸性質によって決定されることが知られている。しかし, これまでの規則化は自立語を中心とした, 各品詞個別の記述が多く, これらを工学的に機械上で実現するためには多くの規則と個別テーブルを必要とする。さらに加えて, 多数個(3個以上)の単語結合に関する記述が少なく, そのままでは十分な指定が難しい。本論文では, 単語の結合によって生ずるアクセント核の移動, 生起, 消失の統一的な規則化を行う。以下, 2.では文節内アクセント核位置決定の規則化を目的とした, 付属語アクセント結合, 複合単語アクセント結合の規則化を中心に, アクセント結合規則を述べる。3.では, この規則の妥当性を検証するために行った文節アクセント推定実験の結果を示す。最後に4.で, これらの検討により得られた知見を整理し, 今後の課題について述べる。

2. アクセント結合規則

いくつかの単語が連なって一つのアクセント的まとまりが作られるとき, これをアクセント句と呼ぶ。日本語では多くの場合, アクセント句は高々1つのアクセント核しか持たないため, その核位置によってアクセント型が決まる。核の位置はモーラを単位として語頭から数えられる。「モーラ」は母音もしくは子音+母音からなる単位であって, 撥音, 促音も1モーラを

[†] 電電公社武蔵野電気通信研究所, 武蔵野市
Musashino Electrical Communication Laboratory, N.T.T.,
Musashino-shi, 180 Japan
論文番号: 昭58-論337[D-102]

形成する。従って4モーラの単語ハジッタ(走った), 6モーラの単語テンキヨウ(天気予報)は各々, 2型, 4型アクセントを持つという。また, ツメタイコウチャ(冷たい紅茶)のように核を持たないものは, 0型アクセントと呼ぶ。アクセント規則を論ずる上では, 長音, 重母音, 母音+撥音を一単位として扱うモーラよりも大きな単位である「音節」も重要な役割を果たす。上に示した例では(ハ)(シッ)(タ):3音節, (テン)(キ)(ヨ)(ホウ):4音節, (ツ)(メ)(タイ)(コウ)(チャ):5音節となる。

このようにアクセント句は1つ以上の文節からなり, 句のアクセント型は, それを構成する文節のアクセント型を反映する。ここでは, まず文節内アクセント結合の規則化を図るため, 付属語アクセント結合の分析を行い, それに基づく規則化を示す。次に, 自立語や接辞の結合として, 複合単語のアクセント結合規則の整理を行う。最後に, 文節間のアクセント結合について触れる。

2.1 付属語アクセント結合規則

2.1.1 付属語アクセント結合の分析

自立語に付属語が結合して文節を作るとき, その文節のアクセントは, 自立語のアクセント型, 文法的諸性質(品詞, 活用種類, 活用形), および付属語のアクセント性質によって種々に変化する。これらを支配する規則を知るため, 付属語に焦点を置き, 自立語(名詞, 動詞, 形容詞)と付属語(助詞, 助動詞)が結合した場合の文節アクセント型を分析した。

いま, N_1 モーラ M_1 型アクセントを持つ自立語に N_2

モーラ $\tilde{M}_2$ 結合アクセント価を持つ付属語が結合して, N_c モーラ M_c 型アクセントを持つ文節ができる場合を考える。ここで, 付属語の結合アクセント価 $\tilde{M}_2$ は結合時に現れる核の位置を示す。従って, 例えば付属語が結合アクセント価0を持つということは, その付属語を単独に発声する際に0型(平板型)になるということではなく, 結合時にその付属語の直前に核が生ずることを意味する。多くの場合, $\tilde{M}_2$ は付属語を単独に発声した場合のアクセント型 M_2 と一致するが, 短い助詞等では必ずしも一致しない。この $\tilde{M}_2$ の導入が本規則化の一つのポイントであり, この表現に基づいて文節アクセント型 M_c と $M_1, \tilde{M}_2$ の関係を調べると, どの場合にも共通した次の性質があることが判明した。

- ◎文節アクセント型 M_c は, 自立語のアクセント型 M_1 と付属語固有のアクセント結合様式ならびに結合アクセント価によって規則的に決定される。
- ◎単語結合に伴うアクセント核位置の移動($M_1 \rightarrow M_c$)傾向を決定するのは, 自立語のアクセント核の有無と付属語固有のアクセント結合様式である。
- ◎アクセント核の移動($M_1 \rightarrow M_c$)の傾向は, 構成単語の品詞情報を用いずに, ほとんど説明できる。

この分析で明らかになった付属語のアクセント結合様式を整理すると, 表1に挙げる5種類に分類することができる。これらの結合様式(1)~(4)-(2)は次の性質を持つ。

- (1) 自立語のアクセント核がそのまま文節のアクセント核となる結合
- (2) アクセント核を持たない自立語に対し, 付属語のアクセント核が現れる結合
- (3) アクセント核を有する自立語に対してのみ, 付属語のアクセント核が現れる結合
- (4)-(1) 自立語のアクセント核の有無に無関係に付属語のアクセント核が現れる結合
- (4)-(2) 自立語のアクセント核の有無に無関係に文節のアクセント核が消失する結合

これらの結合の実際を示すため, アクセント核を持たないもの(ワラウ「笑う」)と持つもの(アルク「歩く」)を例に, 表1に具体例を示した。以下, このアクセント結合様式と結合アクセント価とをあわせて, アクセント属性と呼ぶことにする。

2.1.2 付属語アクセント結合規則とその適用法

前項の分析が示すように, 自立語と付属語のアクセント結合は, 基本的には自立語のアクセント核の有無と付属語のアクセント結合様式によって制御される。

表1 付属語アクセント結合様式

(N_1 モーラ M_1 型 + N_2 モーラ $\tilde{M}_2$ 価 $\rightarrow N_c$ モーラ M_c 型)

アクセント 結合様式	結合後の文節 アクセント型 M_c	具体例
(1) 従属型	M_1	ワラウ [〃] ホド アル [〃] イタ
(2) 不完全支配型	$N_1 + \tilde{M}_2$ ($M_1 = 0$) M_1 ($M_1 \neq 0$)	ワラ [〃] ッタ [〃] リ アル [〃] クヨウダ
(3) 融合型	M_1 ($M_1 = 0$) $N_1 + \tilde{M}_2$ ($M_1 \neq 0$)	ワイナ [〃] ガラ アル [〃] カセル
(4)-(1) 支配型	$N_1 + \tilde{M}_2$	ワラウ [〃] マイ アル [〃] キマス
(4)-(2) 支配型 (平板化型)	0	ワラウ [〃] ダケ アル [〃] クダケ

また、その核の位置は、自立語のアクセント型 M_1 、モーラ長 N_1 、および付属語の結合アクセント価 $\tilde{M}_2$ によって、表1に示す値で与えられる。さらに、これらの結合規則適用にあたっては、以下に示す運用規則が用いられる。

(i) 巡回的適用則

付属語が複数個連なった場合、原則として、アクセント結合規則は左から巡回的に適用される。従って、文節「食べられませんから」を例にとって結合規則の適用法を示すと、以下のようになる。

$\overline{\text{タベル}} + \overline{\text{ラレル}} \rightarrow \overline{\text{タベラレル}}$
 $\overline{\text{タベラレル}} + \overline{\text{マス}} \rightarrow \overline{\text{タベラレマス}}$
 $\overline{\text{タベラレマス}} + \overline{\text{ン}}$ $\rightarrow \overline{\text{タベラレマセン}}$
 $\overline{\text{タベラレマセン}} + \overline{\text{カラ}}$ $\rightarrow \overline{\text{タベラレマセンカラ}}$
 ラレル：融合型，結合アクセント価 $\tilde{M}_2=2$
 マス：支配型，" $\tilde{M}_2=1$
 ン：支配型，" $\tilde{M}_2=0$
 カラ：不完全支配型，" $\tilde{M}_2=0$

(ii) 音節内移動規則

従来よく言われているように、アクセント核の移動は音節単位に行われる。このため、結合規則によって核の移動が生じ、2モーラからなる音節に核が与えられる場合、核は音節中の先行モーラに与えられる。従って、例えば形容詞「優しい」は /ya/, /sa/, /shii/ からなる4モーラ3音節の0型アクセント単語で、不完全支配型、結合アクセント価0を持つ「かしら」が結合すると、核の移動は次のようになる。

$\overline{\text{ヤサシイ}} + \overline{\text{カシラ}} \rightarrow \overline{\text{ヤサシイカシラ}}$
 $\rightarrow \overline{\text{ヤサシイカシラ}}$

(iii) 無声化に伴う移動規則

無声化した母音にアクセント核が移動する場合、原則として、アクセント核は1つ左へ移動する。例えば次のような移動が行われる。

$\overline{\text{ハタラク}} (\text{「働く」}) + \overline{\text{カ}} (\text{不完全支配型, } \tilde{M}_2=0)$
 $\rightarrow \overline{\text{ハタラクカ}}$

(iv) 一段活用動詞処理規則(終止形基本の場合のみ)

動詞のアクセント型の基本を終止形に置くか連用形に置くかは、いろいろと議論のあるところであるが⁽³⁾、結合規則適用上で具体的に問題となつて現れるのは、一段活用の場合である。基本を終止形に置いて考えると、一段活用の未然形、連用形のモーラ数は1つ減少し、アクセント核も左へ1つ移動する。例えば、下一段活用の動詞シラベル(「調べる」)の連用形に従属型の助動詞「た」が結合する場合、

$\overline{\text{シラベル}} + \overline{\text{タ}} \rightarrow \overline{\text{シラベタ}}$ となる。

一方、基本を連用形に置いた場合、上記の規則は不要となり、終止形の活用語尾「る」を融合型で結合アクセント価0を持つ付属語と考えればよい。すなわち前例のシラベル、シラベタは次の結合と考える。

$\overline{\text{シラベ}} + \overline{\text{ル}} (\text{融合型, } \tilde{M}_2=0) \rightarrow \overline{\text{シラベル}}$
 $\overline{\text{シラベ}} + \overline{\text{タ}} (\text{従属型}) \rightarrow \overline{\text{シラベタ}}$

従って本規則は動詞の連用形をアクセント型の基本にとれば不要となる。

(v) 先行品詞による結合アクセント価選択則

本節のはじめで述べたように、結合後の核の位置は M_1 、 N_1 および付属語の結合アクセント価 $\tilde{M}_2$ によって与えられるが、不完全支配型の付属語の一部には $\tilde{M}_2$ の値が自立語の品詞によって異なるものが存在する。これは例えば、

$\overline{\text{サクラ}} + \overline{\text{ナド}} \rightarrow \overline{\text{サクラナド}}$ (桜など)
 $\overline{\text{アソブ}} + \overline{\text{ナド}} \rightarrow \overline{\text{アソブナド}}$ (遊ぶなど)

のように、体言接続と用言接続では核の位置が異なるため、先行品詞に応じた $\tilde{M}_2$ を与える。

2.2 複合単語のアクセント結合規則

自立語や接辞が結合して複合単語を作る場合、その構成単語により複合単語のアクセント型を与える規則については、これまで比較的良く整理されている^{(4),(6)}。ここでは、前節の分析に従い、後続構成単語のアクセント結合様式と結合アクセント価に着目した規則化を図る。

(i) アクセント結合様式

自立語や接辞の結合の場合は付属語結合の場合とは異なり、文節のアクセント型は後続単語の性質によってほとんど決定される。すなわち、2.1.1で示した結合様式のうち(1)~(3)のタイプは複合単語ではほとんどみられず、(4)の支配型が普通である。従って、複合単語の場合、アクセントの差異は主に次に示す結合アクセント価に起因する。

(ii) 結合アクセント価

後続構成単語の自立性が強い場合、結合アクセント価 $\tilde{M}_2$ は単独発声時のアクセント型 M_2 とほとんど一致する。後続単語が平板型、尾高型の名詞である場合はこの例外であり、後続名詞の先頭のモーラに核が生ずる。このため、 $\tilde{M}_2=1$ の場合と同様な振舞いをする。以下、取扱いの一貫性を保つため、この場合は $\tilde{M}_2=1$ として取扱い、特に区別しない。また、動詞、形容詞は取り得るアクセント型がもともと少ないのに加え、複合した場合、どちらも最終モーラ直前のモーラにアクセ

表2 複合名詞におけるアクセント結合の分類(後続名詞の結合アクセント価 $\tilde{M}_2$ による相違)
(N_1 モーラ M_1 型 + N_2 モーラ $\tilde{M}_2$ 結合アクセント価 $\rightarrow N_c$ モーラ M_c 型)

結合の種類	アクセント結合様式	結合アクセント価 $\tilde{M}_2$	複合名詞のアクセント型 M_c	後続名詞の性質	該当する(後続)単語例	結合例
(a) 自立語結合保存型	(4) - (1)	M_2	$N_1 + M_2$	$N_2 \geq 2$ $M_2 \neq 0, (N_2 - 1)^\#, N_2$	(頭高型) 名詞 (中高型)	タイショウサヨウソ (対称作用素)
(b) 自立語結合生起型	(4) - (1)	1	$N_1 + 1$	$N_2 \geq 2$ $M_2 = 0, (N_2 - 1)^\#, N_2$	(平板型) 名詞 (尾高型)	タイショウイドウ (対称移動)
(c) 接辞結合標準型	(4) - (1)	0	N_1	$N_2 \leq 2$	点, 権, 会	タイショウテン (対称点)
(d) 接辞結合平板化型	(4) - (2)	*	0	$N_2 \leq 2$	的, 性, 語	タイショウセイ (対称性)

後続名詞の最終音節が2モーラからなる場合

セント核が生ずることが普通であり, 規則は単純である。従って問題となるのは名詞であり, 特に2モーラ以下の名詞は接辞的性格を持つため, 結合アクセント価 $\tilde{M}_2$ は種々の値を取りうる。このため, 複合名詞におけるアクセント結合について整理し, 結果を表2に示す。

この表が示すように, 結合アクセント価 $\tilde{M}_2$ は単独発声時のアクセント型 M_2 およびモーラ長 N_2 によって(a)~(d)の4種類にわかれる。(a)~(d)は次の性質を持つ名詞, 接辞からなる。

- (a) 最終音節以外にアクセント核を持つ2モーラ以上の名詞
- (b) アクセント核を持たない。もしくは最終音節にアクセント核を持つ。2モーラ以上の名詞
- (c), (d) 2モーラ以下の名詞, 接辞

この分類からも判るように, 2モーラの名詞は(a)~(d)のいずれにも属す可能性があり, どれに属すかは語の自立性, 接辞性の性質等によって決まる。また, アクセント核の位置は音節単位と関連しており, 例えば「入学案内」(ニュウガク+アンナイ $\rightarrow$ ニュウガクアンナイ)に見られるように, 後続名詞「案内」は中高型であるが, 音節的には最終にアクセント核があり, (b)に属すため, このような核が生ずる。

以上述べたように, 複合単語においても付属語の場合と同様に, 後続構成単語のアクセント属性により, アクセント結合の規則化が図れる。また, この結合規則適用に際しては, 前節で述べた適用則のうち, (i) 巡回的適用則, (ii) 音節内移動規則, (iii) 無声化に伴う移動規則がそのまま適用される。

2.3 文節間アクセント結合規則⁽⁷⁾

単語と単語の結合により生ずる文節内のアクセントは, 2.1, 2.2で示したように, 後続単語が決定権を持ち, 基本的には先行単語のアクセント型と後続単語のアクセント属性によって決定される。

一方, 文節が結合して一つのアクセント句を構成する場合は逆に, 先行文節が優先権を持つ。すなわち, 先行文節のアクセント核がない場合のみ, 後続文節のアクセント核が残る。この結合規則を表3に具体例と共に示す。これらの結合規則の適用にあたっては, 文節内アクセント結合の場合同様, 2.1.2で示した(i)~(iii)の規則が用いられる。また, 文節間アクセント結合規則は文節に対して適用されるものであり, 本規則適用前に, 前節までに示した文節内アクセント結合が行われねばならない。

表3 文節間アクセント結合の分類(N_1 モーラ M_1 型 + N_2 モーラ M_2 型 $\rightarrow M_c$ モーラ M_c 型)

アクセント型			結合例
先行文節	後続文節	句	
0	0	0	ツメタイ+コウチャ $\rightarrow$ ツメタイコウチャ(冷たい紅茶)
0	$M_2 (\neq 0)$	$N_1 + M_2$	ヒトリダケ+ノコル $\rightarrow$ ヒトリダケノコル(一人だけ残る)
$M_1 (\neq 0)$	0	M_1	カワイイ+コイヌ $\rightarrow$ カワイイコイヌ(かわいい小犬)
$M_1 (\neq 0)$	$M_2 (\neq 0)$	M_1	アオイ+ミカン $\rightarrow$ アオイミカン(青いミカン)

表4 文節データ・ベース内訳

先行する自立語	サンプル数	文 節 例
クラベル (比べる)	761	クラベルヨウデスゲレド
アルク (歩く)	854	アルカレタ
アヤシイ (怪しい)	507	アヤシカッタヨウニ
ヨワイ (弱い)	507	ヨワイソウデス
コドモ (子供)	408	コドモソガ
アサガオ (朝顔)	408	アサガオクライデ
総 計	3445	

3. 文節アクセント推定実験

2.で述べたアクセント結合規則の有効性を確認すると共に、問題点を知るため、文節アクセントの推定実験を行った。付属語連鎖を含む文節、複合単語それぞれについて推定実験を示すとともに、誤推定をしたサンプルの要因分析を行う。

3.1 付属語連鎖を含む文節のアクセント推定実験

3.1.1 データ・ベース

新聞の語彙調査により得られた助動詞、助詞の連接表⁹⁾に記載されているデータをもとに、会話体表現、文語的表現等を除いて約1200サンプルを得る。これらに対し、表4に示す6個の自立語の中から先行接続可能な語を選び、活用するものは適宜活用させて文節を作成した。なお、これら6語は、動詞、形容詞、名詞それぞれにつき、活用種類、アクセント核の有無等が異なる2語を選んでいる。この結果、表4に示すように計3445文節を得た。これを基本データとする。

次に、これらの文節に対して自動品詞分解¹⁰⁾を施して構成単語の同定を行うと共に、品詞、活用種類、活用形、アクセント属性等を与える。また、この同定に誤りがあれば個別に修正した。

さらに、標準語を話す東京出身の成人男性1名と、標準語を話す成人女性(ナレーター)1名により、作成した各文節にアクセント位置を与えた。このアクセント付与に際して、複数のアクセント型を持つものはそれら全てを列挙すると共に、一文節中に同時に複数個の核を持つものに対しても、核の生起が明確なものはそれら全てを列挙した。

3.1.2 付属語連鎖を含む文節のアクセント性質

アクセント推定実験を行う前に、これらの文節のアクセント型の分布を調べた。結果を各自立語ごとに図1に示す。この分布は単語単独の場合の分布¹⁰⁾の傾向

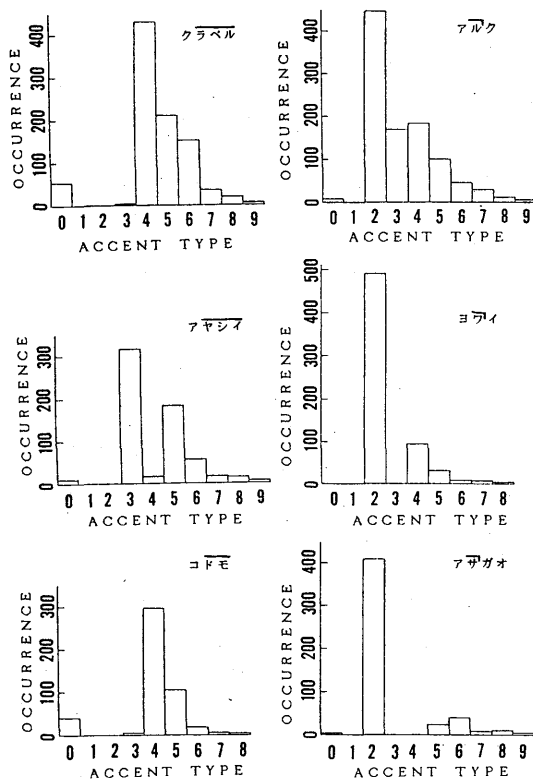


図1 付属語連鎖を含む文節のアクセント型分布

Fig. 1-Accent type distributions of stem-affix concatenations.

とは異なっており、次のような傾向があることがうかがえる。

◎アクセント核を有する自立語は、付属語が連なっても、多くの場合、自立語固有の核は消失しない。また、核の移動が生ずるのは主に動詞で、形容詞、名詞はほとんど移動しない。

◎アクセント核のない自立語に付属語が連なる場合にはほとんど、アクセント核が生ずる。また、その核は語幹の次のモーラ付近に多く生ずる。このため、単語単独の場合と比べ、0型が非常に少ない。

3.1.3 推定実験

3.1.1で示した付属語連鎖を含む3445文節に対し、2.で示した規則によってアクセント型の推定を行った。この推定によって得られる文節のアクセント型の正誤判定は、3.1.1で示したデータ・ベースの値を正解として行った。従って、一文節中に複数個のアクセント核を持つ文節は、それらの核のどれか一つが指定され

表5 付属語連鎖を含む文節のアクセント推定実験結果

自立語	構成品詞数	2	3	4	5	6以上	正解率(%)
比べる	誤り/総数	0/92	0/330	4/256	1/71	2/12	99.1
歩く	誤り/総数	0/114	1/349	3/277	0/93	0/21	99.5
怪しい	誤り/総数	0/68	1/232	4/160	1/39	3/8	98.2
弱い	誤り/総数	0/64	3/235	12/160	7/40	2/8	95.3
子供	誤り/総数	0/59	0/207	1/119	1/19	0/4	99.5
朝顔	誤り/総数	0/59	2/207	9/119	1/19	0/4	97.1
正解率(%)		100.0	99.6	97.0	96.1	87.7	98.3

れば正解となる。

推定実験結果を、各自立語ごとに、文節の構成品詞数別に求め、表5に示す。表に見られるように、各自立語はいずれも95%以上の正解率を示し、全体では、98.3%(3445文節中、誤り58文節)の正解率を得た。この結果、付属語アクセント結合規則は良好な推定を行うことが示され、その有効性が確認された。

3.1.4 アクセント推定誤りの分析

推定誤りを生じた58文節を分析した結果、次の傾向がみられた。

◎文節を構成する品詞数の増加に伴い、正解率は減少する(表5の最下段の正解率を参照)。

◎誤推定アクセント位置はすべて、正しいアクセント位置の後方にある。

◎誤推定アクセント位置のほとんどが、生起の不確定な副次アクセント位置と一致する。

例：子供なのだろう(正解)，子供なのだろう(推定)

◎誤推定となる文節は特定の付属語連鎖を含む。

(たろう，だろう，でしょう，ない)

これらの推定誤りの傾向から、その原因として次の2点が考えられる。

①アクセント結合規則の適用法(2.1.2(i)巡回的適用則)，例：(((ヨワ+カッ)+タ)+デシヨ)+ウ→ヨワカッタデシヨウ(推定)，ヨワカッタデシヨウ(正解)

②付属語の分割単位(例えば学校文法では、

「たろう」=「たろ」+「う」と分割され、「う」のアクセント属性が強く反映され、誤りとなる)。

これら2つの原因は互いに独立したものではなく、共に、単語の基本単位の設定およびそれらの単位による文節の構造等の問題から派生している。

3.2 複合単語のアクセント推定実験

3.2.1 データ・ベース

国語辞典¹⁰⁾に収録されている子見出し7516サン

ルをもとに、アクセント型が明記されている複合単語4877サンプルを選んだ。これを基本データとする。

付属語連鎖を含む文節の場合と同様、これらの文節を自動品詞分解して構成単語の同定を行い、その同定に誤りがあれば個別に修正した。

このデータ・ベースの内訳は、品詞別にみると複合名詞(4735語)が全体の97%を占め、複合動詞(88語)、複合形容詞(46語)は少ない。また、これらを構成単語長別にみると、2個の品詞から構成されるサンプルが全体の95%以上を占め、3個からなるものとあわせると全体の99%以上となる。さらに構成単語長別にみると、2～5モーラの組合せがほとんどであり、(3～4モーラ)+(2～4モーラ)の組合せで全体の約2/3を占める。

3.2.2 推定実験

2.2で示したように、複合名詞のアクセント決定には、後続単語の結合アクセント価が重要な役割を果たす。前述したように、2モーラ以下の名詞については単独発声時のアクセント型がそのまま結合アクセント価となるとは限らないため、個別に与える必要がある。このため、推定実験に先だち、2モーラ以下の後続名詞を抜き出し、それらの結合アクセント価を与えた。これらのうち接辞結合を生ずるものは、異なりで544語あり、標準型350語、平板化型194語からなっている。

以上の準備の上で、複合単語4877サンプルに対してアクセント推定実験を行った。正解としては、辞書に記載されていたアクセント型および他のアクセント辞典^{7)~9)}に記載されていたアクセント型を用い、これらのどれかに一致すれば正解とした。結果を表6に示す。表が示すように、正解率は、複合名詞：95.4%、複合動詞：97.7%、複合形容詞：100.0%となり、付属語連鎖を含む文節同様、複合単語においても、規則

表6 複合単語のアクセント推定実験結果

品 詞	構 成 品 詞 内 容	サンプル 総 数	誤 推 定 語 数		正 解 率 (%)
			誤り①*	誤り②*	
複 合 名 詞	名詞＋名詞	4 0 5 1	1 1 6	7 7	9 5 . 4
	動詞連用形＋名詞	2 0 9	0		
	名詞＋名詞＋名詞	1 2 8	5		
	名詞＋動詞連用形	2 4 8	2	1 4	
	動詞連用形＋動詞連用形	4 3	0		
	そ の 他	5 6	4		
	総 合	4 7 3 5	1 2 7		
複 合 動 詞	名詞＋動詞	3 2	0	2	9 7 . 7
	動詞連用形＋動詞	5 6	2		
	総 合	8 8	2		
複 合 形 容 詞	名詞＋形容詞	3 2	0	0	1 0 0 . 0
	動詞連用形＋形容詞	1 3	0		
	そ の 他	1	0		
	総 合	4 6	0		
その他	(複合副詞)	8	2	2	7 5

* アクセント結合規則に起因する誤り

** 結合アクセント価指定に起因する誤り

の妥当性が確認できた。

3.2.3 アクセント推定誤りの分析

4877サンプル中、誤推定は222サンプルあり、その原因を品詞別に考察する。

(a) 複合名詞

表6に示したように、複合名詞の誤り218語を分類すると、①結合規則に起因するもの127語と、②後続名詞の結合アクセント価に起因する誤り91語に大別される。

①複合単語アクセント結合規則に起因する誤り

この誤りに属す複合語は、通常、アクセント結合に重要な役割を果たす後続単語のアクセント性質が現れず、先行単語のアクセント性質がそのまま現れる。さらに細かく分類すると次のように分かれる。

①-(i) 先行単語アクセント残存型(81語)

先行単語の核がそのまま複合単語のアクセント核となる。多くは1型アクセントを持つ。

例: ギシカイセイ (起死回生)

オチヨウメチヨウ (雄蝶雌蝶)

①-(ii) 平板型(41語)

先行単語同様、複合語も平板型となる。

例: ギョクセギコンコウ (玉石混淆)

ゼンシンゼンレイ (全身全霊)

①-(iii) その他(5語)

ナйкаク・カンボウチヨウカン (内閣官房長官)

ハカイカツドウボウシホウ (破壊活動防止法)

デンバボウエンキョウ (電波望遠鏡)

デンシケンビキョウ (電子顕微鏡)

ヒガシニホン (東日本)

これらの例に示されるように、(i),(ii)に属す複合語は独立した一単語の色彩が強いものか対句をなすものであり、通常の複合語とは性質が異なる。また(iii)に挙げたものは複合語の語構成と関連している。

例えば、「電波望遠鏡」を構造上からみれば、(電波+(望遠+鏡))の形をしており、結合規則はこの順序に従って適用されるので、デンバ+ボウエンキョウ→デンバボウエンキョウと誤る⁽⁵⁾。

② 後続単語の結合アクセント価に起因する誤り

各後続単語には1つの結合アクセント価が与えられているが、接辞的結合をするものの中には2種(標準型、平板化型)の型をもち、用法、意味によって使い分けられるものが存在する。

例: 「式」 (カガクシキ (化学式))

カリョクシキ (火力式)

「目」 (ハヤリメ (はやり目))

オサエメ (押え目)

「猫」 (ベルジャネコ (ベルシャ猫))

ミケネコ (三毛猫)

(b) 複合動詞、複合形容詞、その他

これらで問題となったのは次の4語であり、必ずしも誤りとはいいい切れないものも含まれている(前2例): 恐れ入る、驚き入る、ずるずるべったり、ずんぐりむっくり。これらの誤りのうち、「入る」は補助用言的使用のため、動詞としてより付属語的な結合となる。また、他の2例に関しても原因分析は行えるが、サンプル数も少なく、十分な議論ができないので、それらの考察は今後の検討に譲りたい。

4. む す び

日本語テキストからの音声合成等に用いられる韻律情報の生成規則化の第一歩として、アクセント句を構成する単語の性質からアクセント核位置を決定する規則の統一化を図った。このため、単語間アクセント結合において、後続単語のアクセント属性として、結合アクセント価、アクセント結合様式を提案した。これにより、従来個別的な記述がされていた付属語のアクセント結合が整理されると共に、複合単語についても

同じ枠組で扱えるようになった。また、これらの規則の有効性を示すため、付属語連鎖を含む文節3445サンプル、複合単語4877語を用いたアクセント推定実験を行った。この結果、それぞれ98.3%、95.4%の正解率を得、規則の妥当性が示された。

さらに誤推定原因の分析により、アクセント決定には、アクセント句を構成する単位、それらによる句の構造、語の用法や意味等と深い関連があることが判明した。今後は、これらアクセント句を構成する単位の規定、句構成とアクセント結合の適用法等の関係を明らかにしてゆくと共に、ここで取り上げなかった接頭辞、補助用言等についても本規則化を図ることが一つの課題である。さらに、長い文節における主アクセントと副次アクセントの指定方法、語構成とビッチパタンの関係等、音響面からの分析もあわせて行い、韻律制御規則化を進めてゆきたい。

謝辞 本研究を進めるのにあたり種々御指導いただいた畔柳基礎部長、橋本画像部長、ならびに板倉第四研究室長に感謝します。また、自動品詞分解に協力いただいた小暮社員に感謝します。

文 献

- (1) IEEE ICASSP'82 pp.1589-1611 (Text to Speech System に関する数カ国の報告)。
- (2) 佐藤, 匂坂, 小暮, 嵯峨山: “日本語テキストからの音声合成”, 昭57通信総全大, S6-3。
- (3) McCawley, J.D.: “The Phonological Component of a Grammar of Japanese”, Mouton(1968)。

- (4) 和田 実: “辞のアクセント”, 論集・日本語研究2, アクセント, 有精堂(昭55)。
 - (5) 早田輝洋: “放送での発音とアクセント”, NHK総合放送文化研究所編, 放送用語論, pp.153-199(昭50)。
 - (6) 早田輝洋: “東京方言の音韻化規則”, 言語研究, 49号, pp.55-69(1966)。
 - (7) 秋永一枝: “共通語のアクセント”, 日本放送協会編, 日本語アクセント辞典, 所載
 - (8) 金田一春彦監修: “明解日本語アクセント辞典”, 三省堂(昭42)。
 - (9) 平山輝男編: “全国アクセント辞典”, 東京堂(昭37)。
 - (10) 橋本新一郎: “日本語単語アクセントの諸性質”, 信学論(D), 56-D, 11, pp.654-661(昭48-11)。
 - (11) 佐藤, 比企: “代名語を材料とした日本語アクセントパターンに関する一考察”, 信学誌(D), 57-D, 8, pp.471-478(昭49-08)。
 - (12) 匂坂, 佐藤: “複合単語のアクセント規則の検討”, 日本音響学会音声研究会資料, S81-68(1982-01)。
 - (13) 匂坂, 佐藤: “付属語連鎖におけるアクセント結合の分析”, 日本音響学会音声研究会資料, S82-27(1982-07)。
 - (14) 壁谷, 石川: “文音声変換におけるアクセント付与規則の検討”, 情報処理学会自然言語処理研究会資料, 31-8(1982-05)。
 - (15) 国立国語研究所: “電子計算機による新聞の語彙調査Ⅱ”, 秀英出版(昭47)。
 - (16) 小暮, 嵯峨山, 佐藤: “音声合成のための日本語テキスト解析”, 日本音響学会講演論文集, 1-5-18(昭57春季)。
 - (17) 金田一京助他編: “新明解国語辞典第二版”, 三省堂(昭49)。
- (昭和57年11月12日受付, 58年1月31日再受付)