

# 修士論文

日本語音声合成のためのアクセント  
結合規則の改善とデータベースに  
基づく統計的アクセント処理

2007年2月2日 提出

指導教員 峯松 信明 助教授

東京大学大学院 情報理工学系研究科 電子情報学専攻

56417 黒岩 龍

---

# 内容梗概

日本語の文における各単語のアクセントは、各単語を単独で発声する場合のアクセントとは異なるものになる。日本語テキスト音声合成システムでは、形態素解析結果に含まれる各単語の単独発声アクセントから文中アクセントを導き出す処理が必要となる。この処理を統一的な規則として表現する研究がなされており、テキスト音声合成システムとしての実装もなされているが、誤った出力もしばしば見られ、改善を必要とする。

本研究では、規則の改善と統計的手法の導入の2つの観点から、より精度の高いアクセント処理の実現を目指した。規則の改善には、まず、規則による処理が持つ意味を的確に捉え、整理しなおすことによって規則の単純化を実現し、さらに用言等の活用による影響を規則に取り入れることで精度の向上を実現した。また、統計的手法の導入にあたっては、学習に必須となる正解ラベル付きデータが存在しなかったため、まず、アクセントデータベースを作成するところから行なった。そして、作成したデータベースを使用して統計的学習を行ない、規則から得られる知識を取り入れるなどの工夫をした結果、規則による処理を上回る高い正答率が得ることができた。

# 目次

第1章	序論	1
第2章	日本語アクセント結合処理とその現状	3
2.1	アクセントの基礎	3
2.1.1	アクセント句とアクセント核	3
2.1.2	アクセント型	3
2.2	句坂らによるアクセント結合規則	4
2.2.1	結合規則の概要	5
2.2.2	付属語アクセント結合規則	5
2.2.3	複合単語アクセント結合規則	6
2.2.4	接頭辞アクセント結合規則	7
2.2.5	文節間アクセント結合規則	8
2.2.6	アクセント結合規則適用則	8
2.3	音声合成システムにおける句坂らの規則の実装	9
2.3.1	GalateaTalkの構成	9
2.3.2	Gtalk,UniDicの品詞体系	9
2.3.3	Gtalk,UniDicにおける先行品詞別属性選択	10
2.4	田中による付属語アクセント	10
2.5	長野による確率モデルを用いた読み及びアクセント推定	11
第3章	アクセント結合規則の改良	14
3.1	基本となる考え方	14
3.1.1	句坂らの「様式」と田中の「式」の比較	14
3.1.2	様式間の共通性への着目	14
3.1.3	活用の考慮	15
3.1.4	先行品詞別属性選択の是非	16
3.2	活用による影響の検討	16
3.2.1	動詞の活用	16
3.2.2	形容詞の活用	19
3.2.3	助動詞の活用	19
3.3	活用を考慮したアクセント推定実験	21
3.4	その他の課題点	21
第4章	単独の日本語話者によるアクセントデータベースの作成	23
4.1	アクセントデータベース作成の目的	23
4.2	アクセントデータベースの仕様	23

4.2.1	ラベリング内容	23
4.2.2	ラベリング対象とする言語事象の定義	23
4.2.3	ラベラを1人とする理由及び単独発声ラベリングの目的	25
4.2.4	使用した文セット	26
4.2.5	形態素解析	27
4.2.6	読み仮名の修正	28
4.2.7	ラベリング従事者の選定	28
4.2.8	実際のラベリング作業の詳細	28
4.2.9	ラベル検査作業・再ラベリング作業	32
4.3	完成状況と分析	32
4.3.1	完成状況	32
4.3.2	分析	32
<b>第5章</b>	<b>統計的手法を用いたアクセント処理</b>	<b>34</b>
5.1	統計的手法の導入	34
5.2	条件付確率場	34
5.2.1	条件付確率場	34
5.2.2	CRF++	35
5.3	条件付確率場を用いたアクセント処理の学習・推定	36
5.3.1	学習・推定の内容	36
5.3.2	単独発声アクセント型を与えない場合における文中アクセント型の学習・推定	38
5.3.3	単独発声アクセント型を与える場合における文中アクセント型の直接学習・推定	39
5.3.4	規則に基づいた文中アクセント型の推定	39
5.3.5	考察	41
5.4	アクセント結合規則を参考とした学習・推定の工夫	42
5.4.1	単独発声アクセント型から文中アクセント型への変化についての学習・推定	42
5.4.2	隣接形態素組み合わせ素性を用いた学習・推定	43
5.4.3	アクセント結合規則に適合させた学習・推定	44
5.4.4	不正解に分類された形態素の許容度に関する調査	47
5.4.5	考察	48
<b>第6章</b>	<b>結論</b>	<b>49</b>
6.1	本研究のまとめ	49
6.2	今後の課題	49
	<b>参考文献</b>	<b>52</b>
	<b>発表文献</b>	<b>54</b>
	<b>付録A ラベラに配布した説明文書</b>	<b>55</b>

# 目次

1.1	音声合成における処理の流れ	1
2.1	アクセント型の種類（4モーラ語の場合）	4
3.1	活用によるアクセント変化の考慮	16
4.1	複数人でラベリングをすることの弊害	26
4.2	ラベリング作業用プログラムによる文発声ラベリング	29
4.3	形態素単独発声ラベリング	30
4.4	ラベリング作業のフローチャート（文データの流れ）	31
5.1	観測データ（ $x$ ）と出力ラベル（ $y$ ）	34
5.2	観測素性	35
5.3	遷移素性	35
5.4	アクセント結合	36
5.5	アクセント結合を個々の形態素のアクセント変化と捉える場合	36
5.6	形態素解析結果と対応付けたアクセントラベル	37
5.7	許容度判定プログラム	47

# 表 目 次

2.1	付属語アクセント結合様式 . . . . .	5
2.2	複合名詞アクセント結合の分類 . . . . .	6
2.3	接頭辞アクセント結合の分類 . . . . .	7
2.4	文節間アクセント結合 . . . . .	8
2.5	田中による付属語アクセントの「式」 . . . . .	11
2.6	複合名詞「京都タワー」における「京都」及び「タワー」のアクセント . . . . .	12
2.7	複合名詞「京都タワーホテル」における「京都」「タワー」及び「ホテル」の アクセント . . . . .	12
2.8	長野らによるアクセント処理の精度 . . . . .	12
3.1	匂坂らの「様式」と田中の「式」の比較 . . . . .	14
3.2	書き換えた付属語結合規則 . . . . .	15
3.3	単純化した付属語結合規則 . . . . .	15
3.4	動詞-五段活用（歩く） . . . . .	17
3.5	動詞-五段活用（笑う） . . . . .	17
3.6	動詞-一段活用（見る） . . . . .	17
3.7	動詞-一段活用（生きる） . . . . .	17
3.8	動詞-一段活用（居る） . . . . .	17
3.9	動詞-一段活用（借りる） . . . . .	17
3.10	動詞-カ変活用（来る） . . . . .	18
3.11	動詞-サ変活用（する（追求する, . . . . .）） . . . . .	18
3.12	動詞-サ変活用（発する（感ずる, . . . . .）） . . . . .	18
3.13	動詞-サ変活用（利する（信ずる, . . . . .）） . . . . .	18
3.14	形容詞-現代形容詞活用（強い） . . . . .	20
3.15	形容詞-現代形容詞活用（鋭い） . . . . .	20
3.16	形容詞-現代形容詞活用（赤い） . . . . .	20
3.17	形容詞-現代形容詞活用（危ない） . . . . .	20
3.18	助動詞-ダ活用（だ） . . . . .	20
3.19	助動詞-デス活用（です） . . . . .	20
4.1	形態素単独発声ラベリングでの形態素の提示方法 . . . . .	31
4.2	アクセント句を構成する形態素の数 . . . . .	33
4.3	アクセント句を構成する品詞列（出現率1%以上） . . . . .	33
5.1	統計処理による正答数（単独型なし, 当該形態素の特徴のみ使用） . . . . .	38
5.2	統計処理による正答数（単独型なし, 前後1形態素までの特徴を使用） . . . . .	38
5.3	統計処理による正答数（単独型なし, 前後2形態素までの特徴を使用） . . . . .	38

5.4	統計処理による正答数（単独型あり, 当該形態素の特徴のみ使用）	40
5.5	統計処理による正答数（単独型あり, 前後 1 形態素までの特徴を使用）	40
5.6	統計処理による正答数（単独型あり, 前後 2 形態素までの特徴を使用）	40
5.7	規則に基づく手法の正答数	40
5.8	相対変化ラベルの学習による正答数（前後 2 形態素までの特徴を使用）	43
5.9	相対変化ラベルの学習による正答数（前後と組み合わせた特徴も加えて使用）	44
5.10	規則を参考とした変化の学習による正答数	46
5.11	許容度調査の結果	48
5.12	規則を参考とした変化の学習による正答数（許容される句も正答とする場合）	48

# 第1章 序論

今日の日常生活において、機械による音声案内は多くの場所で活用されている。公共交通機関での旅客案内やテレホンバンキングでの自動ガイダンスなどは既に馴染み深い。今後情報家電に代表されるようなユビキタスコンピューティングが一般的になれば、音声による自然な出力が一層期待されることと思われる。

現在実用化している音声出力の中には、必要な文章全体が録音されたものを再生するシステムや、ある程度の長さで録音したものを継ぎはぎすることで文を構成するシステムが多数存在するが、これが可能なのは出力文のパターンが少数に限られる場合のみであり、多様な出力が求められる場合には任意のテキストからの音声合成が可能でなくてはならない。しかし、任意のテキストからの音声合成においては、一般的に入力される情報は文字列としての文章のみであり、そこから自然な音声を得るには複雑な仕組みを必要とするため大変難しく、いまだ十分な技術は得られていない。日本語のテキスト音声合成においても例に漏れず、同様といえる。

日本語テキスト音声合成ソフトウェアに漢字仮名混じり文のテキストが入力された場合、一般的には、まず付属する単語辞書を参照してテキスト解析（形態素解析等）を行なうことによって単語境界や各単語の読み（モーラ列もしくは音素列）・基本アクセント型などの情報を得、それを基に音韻処理（無声化などを含む処理）と韻律処理（アクセント・イントネーション・継続長・パワーなどの処理）がそれぞれ行なわれ、最終的な音声波形出力になる [1]。はじめに辞書を用いて得られたアクセントの情報は個々の単語を単独発声したときのものであり、後の韻律処理において、文としてのアクセントを得るための的確な処理が必要となる。

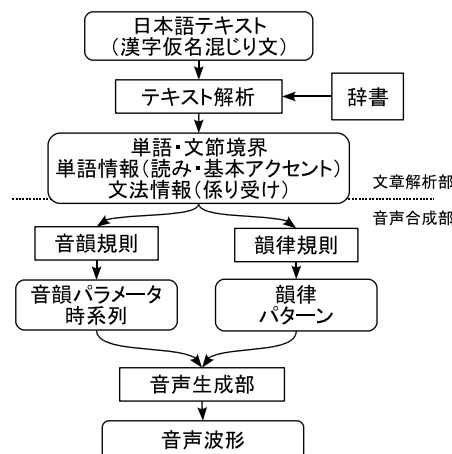


図 1.1: 音声合成における処理の流れ

日本語の韻律に関する研究のためには、この処理の詳細を知ることは重要である。しか



し、過去の言語学・工学のいずれにおける研究成果においても、アクセント処理の網羅的な説明は達成されておらず、いずれの成果を用いてもあらゆる構成を持つ文に対してアクセント処理の内容を説明することは不可能であった。いずれの学問分野においても、アクセント処理の体系的な説明は日本語の韻律研究に関する知識の基盤として必要と考えられるが、それはいまだ不十分である。

日本語テキスト音声合成での利用を想定してアクセント処理を統一的な規則で表現した研究としては、句坂らによるアクセント結合規則の研究が挙げられる ([2, 3])。句坂らによれば、日本語（東京語）のアクセントに関する研究は言語学分野での研究が盛んなわりに工学的な立場での研究は少なかった。アクセント句（いくつかの単語が連なって作られる1つのアクセント的まとまり）のアクセントはそのアクセント句を構成する単語それぞれの性質から決定されることが知られているが、それまでの規則化では品詞ごとの記述に分かれ、工学的に機械上に実現するには煩雑すぎであった。さらに、多数の単語が連なるようなものの記述は少なく、これでは実用的なレベルでのアクセント推定には不十分であった。これを、句坂らはそれぞれの単語や文節などにアクセントに関する属性を定義し、単語が連なることによるアクセントの変化を統一的な規則によって表現し、それに基づくアクセント推定実験では高い正答率を得た。

情報処理技術振興協会（IPA）の Galatea Project で開発された音声合成モジュールの GalateaTalk[4, 5] は、これまでに述べてきたような任意のテキストからの音声合成システムである。形態素解析をおこなった後に行なうアクセント処理については、句坂らによって提案された規則に基づき、具体的な各種属性値の決定 [6] を実装して用いている。

しかし、句坂らの規則は単語の結合の種類についてあらゆる場合を考慮されてはおらず、結果として、GalateaTalk では誤った出力がなされる場合も少なくない。実際に日常に現れるテキストは、必ずしも単純なものばかりとは限らず、そのような多種多様な入力に対し、テキスト音声合成システムは適切なアクセントでの出力をしなければならぬため、句坂らの規則は、実用上において不十分な点があると言える。

本研究では、あらゆるテキストに適切なアクセントを付与するシステムの開発を目的とした。はじめに、第2章では、日本語のアクセントについての基礎的な知識について説明するとともに、句坂らの規則を中心にアクセント処理に関する研究の現状について述べる。第3章では、句坂らの規則に意味づけを行なうことで規則の単純化をし、その単純化された規則に対して、用言等の活用への考慮を組み入れたことによる、規則の精度向上について述べる。第4章以降では、正解ラベルを付与されたデータを用いた統計的学習について述べるが、まず第4章では、統計的学習に用いるためのデータベース作成について説明する。第5章では、作成したデータベースを使用して行なった統計的アクセント学習・推定処理に関して述べる。最後に第6章では、今後の課題についても触れた上で、本研究についてまとめる。

## 第2章 日本語アクセント結合処理とその現状

### 2.1 アクセントの基礎

#### 2.1.1 アクセント句とアクセント核

アクセントと呼ばれる現象は多くの言語に存在するが、言語によって大きな違いがある。アクセントとは、言葉を声で発するとき短い区間内で何らかのメリハリを付けることを表すが、実際の表現形態としては強勢アクセント(ストレスアクセント)・高低アクセント(ピッチアクセント)の2種類に大きく分かれる。

強勢アクセントとは音の強さの変化で表現するアクセントであり、英語などで使われている。なお、強さの変化に伴って高さや継続長にも自ずと変化が現れるが、このことについて、英語などのアクセントは強さのみではなく複数の音響量によって表現されているのだとする意見もある。それに対し、高低アクセントは音の高さの変化によるアクセントであり、日本語におけるアクセントはこの高低アクセントである。

さて、日本語の単語を音韻の面から細かく分けたものに、“モーラ”(“拍”とも呼ばれる)がある。モーラとは母音1つ、または子音と母音を組み合わせたもので、およそカナ1文字に対応する。撥音(ン)や促音(ッ)も1つのモーラとなるが、拗音(ャ・ヨなど)は2文字で1つのモーラを形成する。伸ばした音(オーなど)は2モーラとなる。すなわち、「ロンブン」は4モーラ、「テイシュツ」は4モーラ、「モーラ」は3モーラである。

日本語のアクセントは、このモーラ単位の音の高さの変化として生ずる。実際の会話においては強弱・高低とも多様に变化し、特に感情がこもればいっそう激しいものになるが、文の全体にわたる緩やかな変化(イントネーション)を取り除けば、一定の範囲ごとに、アクセントによる決まった高低変化が現れる。このような、いくつかの単語が連なって作られる1つのアクセント的まとまりのことをアクセント句といい、アクセント句を構成する単語によって当該アクセント句でのアクセントの型が決まる。

東京方言においては、音の高さが下降する箇所のみがアクセントとしての重要な意味を持ち、上昇する箇所はそうでない([12, 13, 14])。また、音の高さの上昇はアクセント句の先頭のみにはしばしば発生するものであり、アクセント句としてのまとまりが始まることを表すのみでしかない。よって、アクセント句は、先頭で音の高さの上昇が発生し、以後は個々の単語が持つアクセント的特徴の表れとしての下降のみが起こる構造を持つ。この下降する箇所、もしくは下降が起こる直前のモーラのことをアクセント核と呼ぶ。

#### 2.1.2 アクセント型

アクセントによって現れる音の高低を抽象化し、音の高さを高・低の2値のみで表すことが可能であるとされることが多く、実際に多くの場合はこれに当てはまる。この考え方に基づき、単語などのアクセントの型を表現するため、「 $\bar{\text{ウ}}$ ミ」などのように高の部分の上に線

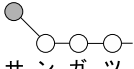
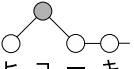
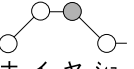
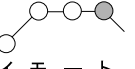
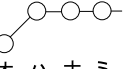
起伏型				平板型
頭高型	中高型		尾高型	φ型
1型	2型	3型	4型	
 サンガツ	 ヒコーキ	 ナイヤシユ	 イモート	 オハナミ

図 2.1: アクセント型の種類 (4 モーラ語の場合)

を引いて記することが多い。

前述の通り、アクセント句の内部では音の高さの上昇は発生しないため、音の高さを高・低の2値のみで表せるものとする場合には、音の高さの下降（高から低に変化する箇所）もアクセント句に高々1つしか存在しないこととなる。つまり、アクセント句の1モーラ目と2モーラ目の間に必ず低から高もしくは高から低に変化し、それ以外の箇所では低から高には変化せず、高から低への変化のみが起こりうることとなる。よって、どの箇所でも高から低に変化するかというアクセント核の位置を示すことでアクセントの型を区別することができる。核が  $M$  モーラ目にある場合、アクセント型は  $M$  型と表現される。核が無いものは一般的には0型と表すが、本論文では核が無いことを明確にするため  $\phi$  型と表すこととする<sup>1</sup>。

この方法で表現すると、例えば5つの単語「三月」「飛行機」「内野手」「妹」「お花見」は、1型: 「サンガツ(三月)」、2型: 「ヒコーキ(飛行機)」、3型: 「ナイヤシユ(内野手)」、4型: 「イモート(妹)」、 $\phi$ 型: 「オハナミ(お花見)」として型を区別することができる。イモートとオハナミは単独での発声時には違いが無いが、後に助詞の「が」を付けると「イモートガ」「オハナミガ」と違いが現れるため、別のアクセント型として扱う。

アクセント核があるものを起伏型、核がないもの( $\phi$ 型)を平板型ということもある。また、起伏型のうち、核が1モーラ目にあるもの(1型)を頭高型、中間のモーラにあるものを中高型、最終モーラにあるものを尾高型とさらに分けることがある。

また、アクセントを考える際に“音節”も重要な役割を果たすことがある。日本語の音節は、モーラに分けて表現したもののうち、長音・重母音・母音+撥音/促音が1単位にまとまったものとしていうことができる。例えば、「ガツ/ショ/レン/メイ」(合唱連盟)は4音節である。

## 2.2 句坂らによるアクセント結合規則

ここでは句坂らによるアクセント結合規則について解説する。なお、[2]中に明記されていないが、品詞体系としてはいわゆる学校文法を想定しているものと思われる。なお、以後では、句坂らによる規則を実装したソフトウェアとして GalateaTalk を用いるが、実装に際して手を加えられているところ ([6, 10, 11]) については、適宜そちらに合わせて説明する。

<sup>1</sup> $\phi$ 型という表現方法の利点については、第3章で触れる。

表 2.1: 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  $\widetilde{M}_2$  価  $N_c$  モーラ  $M_c$  型)

アクセント 結合様式	結合後のアクセント型 $M_c$		具体例
	$M_1 = \phi$ のとき	$M_1 \neq \phi$ のとき	
F1	$M_1$		が (格助詞), た (助動詞)
F2	$N_1 + \widetilde{M}_2$	$M_1$	か (終助詞), です (助動詞)
F3	$M_1$	$N_1 + \widetilde{M}_2$	せる (助動詞)
F4	$N_1 + \widetilde{M}_2$		ます (助動詞)
F5	$\phi$		だけ (副助詞)
F6	$N_1 + \widetilde{M}_{2a}$	$N_1 + \widetilde{M}_{2b}$	ず (助動詞)

### 2.2.1 結合規則の概要

アクセント句は1つ以上の文節によって構成され、句のアクセント型は構成する文節のアクセント型を反映する。また、文節は複数の語からなり、文節のアクセント型はやはりその構成語の性質によって決定されるということができる。

### 2.2.2 付属語アクセント結合規則

自立語 (名詞・動詞・形容詞等) と付属語 (助詞・助動詞) が結合して文節ができるときの規則である。このような結合のできる文節は、自立語のアクセント型や品詞・活用種類・活用形などの文法的諸性質とともに、付属語のアクセントに関する性質が強く影響しており、結合規則は付属語に焦点をおいたものとなっている。

まず、後続の付属語に“結合アクセント価”というものを定義する。結合アクセント価とは、結合した結果としてその語の中にアクセント核が現れる場合のアクセント核の位置を、(結合前の) その語の先頭から数えたものである。例えば、結合アクセント価が2であればその語の2モーラ目が結合後に核となることを意味し、結合アクセント価が0であればその語の直前が核となることを意味する。付属語のアクセント型と結合アクセント価は多くの場合一致するが、短い助詞の場合には一致しないこともある。

$N_1$  モーラで  $M_1$  型のアクセント型を持つ自立語と  $N_2$  モーラで  $\widetilde{M}_2$  価の結合アクセント価を持つ付属語を結合して  $N_c$  モーラで  $M_c$  型の文節ができるとすると、その様子は表 2.1 のように分類され、そのそれぞれを“アクセント結合様式”という。

これには、次のような共通した性質がある。

- 文節アクセント型  $M_c$  は、自立語のアクセント型  $M_1$  と付属語のアクセント結合様式ならびに結合アクセント価  $\widetilde{M}_2$  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 ( $M_1 \rightarrow M_c$ ) の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。

表 2.2: 複合名詞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 ·  $\widetilde{M}_2$  価  $N_c$  モーラ  $M_c$  型 )

種類	後続語の性質	$\widetilde{M}_2$	$M_c$	具体例
C1	$N_2 \geq 2$ $M_2 \neq (N_2 - 1, N_2, \phi)$	$M_2$	$N_1 + \widetilde{M}_2$	手続き (テツツキ), 日間 (ニチカン)
C2	$N_2 \geq 2$ $M_2 = (N_2 - 1, N_2, \phi)$	1	$N_1 + 1$	生活 (セイカツ), 時間 (ジカン)
C3	$N_2 \leq 2$	0	$N_1$	湾 (ウン), 学 (ガク)
C4	$N_2 \leq 2$	*	$\phi$	島 (シマ), 系 (ケイ)
C5	—	*	$M_1$	殿 (ドノ)
C10	—	*	$M_1   M_2$ (双方が残る)	等々 (トウトウ)

そして、アクセント結合様式には以下のような種類がある。

- F1 (従属型) : 自立語のアクセント核がそのまま文節のアクセント核になる。
- F2 (不完全支配型) : アクセント核を持たない自立語に結合した場合にのみ、付属語部分にアクセント核が現れる。
- F3 (融合型) : アクセント核を持つ自立語に結合した場合にのみ付属語部分にアクセント核が現れる。
- F4 (支配型 1) : 自立語のアクセント核の有無に無関係に付属語部分にアクセント核が現れる。
- F5 (支配型 2) : 自立語のアクセント核の有無に無関係にアクセント核が消失する。
- F6 : 自立語のアクセント核の有無に応じて、付属語部分に異なるアクセント核が現れる。

これを用いると、付属語をアクセント核を持つ自立語と持たない自立語にそれぞれ結合させて結果をみれば、その付属語のアクセント結合様式と結合アクセント価を決定することができる。また、付属語についてこれらが分かっているならば、任意の自立語に結合させた結果を規則から導くことができる。

### 2.2.3 複合単語アクセント結合規則

複数の自立語の結合や自立語と接尾辞の結合によって複合単語ができるときの規則である。一部のものを除き、結合の結果には先行する自立語のアクセント核は現れない。すなわち、支配型 (表 2.1 の F4 または F5) になり、後続単語・接辞の結合アクセント価に起因するものとなる。

表 2.3: 接頭辞アクセント結合の分類  
 ( $N_1$  モーラ  $M_1$  型 +  $N_2$  モーラ  $\tilde{M}_2$  価  $N_c$  モーラ  $M_c$  型)

種類	結合後のアクセント型 $M_c$		
	$M_2 = N_2, \phi$ のとき	$M_2 \neq N_2, \phi$ のとき	具体例
P1	$\phi$	$N_1 + M_2$	御 (ゴ)
P2	$N_1 + 1$	$N_1 + M_2$	総 (ソウ)
P13	$M_1$		現 (ゲン)
	$M_1   M_2$		
P4	$N_1 + 1$	$N_1 + M_2$	両 (リウ)
	$M_1$		
	$M_1   M_2$		
P6	$\phi$		再来 (サライ)
P14	$M_1$	$N_1 + M_2$	要 (ヨウ)

後続単語の結合アクセント価は、後続単語が名詞である場合、ほとんどの場合は単独発声時のアクセント型と一致するが、後続名詞にアクセント核が無いまたはアクセント核が最終音節内にあるときは、結合アクセント価は1となる。また、短い名詞(2モーラ以下)の場合は接尾辞としての性格を持つことがあり、この場合は長い名詞(2モーラ以上)の場合とは違い、先行単語の最終モーラに核を生じさせるものと平板化させるものの2種類がある。2モーラの名詞はいずれになることもありうる。他に特殊なものとして、先行する自立語の核を残すもの、韻律的な複合をせず先行する自立語と自身の双方の核が現れるものがある。この様子を表したのが表 2.2 である。

なお、後続単語が動詞・形容詞の場合は、結合後は通常の動詞・形容詞と同様に核を持たないか最終音節の直前に核があるかのいずれかの形にしかならない。後者となるのが普通である。

#### 2.2.4 接頭辞アクセント結合規則

接頭辞と自立語が結合するときの規則である。後続の自立語が動詞・形容詞の時は前述のようになるが、名詞のときは特別な変化が見られる場合がある。接頭辞では、後部要素が尾高型・平板型であるかどうかによって適用される式が区別される。結合の種類は、表 2.3 のように分類される。

P13 および P4 に属する接頭辞の結合した句は、複数のアクセント型を取ることが可能であるが、これらは構文・意味上の違いによって使い分けられる。

表 2.4: 文節間アクセント結合  
 ( $N_1$  モーラ  $M_1$  型文節 +  $N_2$  モーラ  $M_2$  型文節  $N_c$  モーラ  $M_c$  型文節)

アクセント型		
先行文節	後続文節	結合後のアクセント句
$\phi$	$\phi$	$\phi$
$\phi$	$M_2(\neq \phi)$	$N_1 + M_2$
$M_1(\neq \phi)$	$\phi$	$M_1$
$M_1(\neq \phi)$	$M_2(\neq \phi)$	$M_1$

### 2.2.5 文節間アクセント結合規則

アクセント句が複数の文節によって形成される場合の、文節と文節の結合に関する規則である<sup>2</sup>。この場合は、複合名詞とは逆に先行する文節が優先権を持ち、先行する文節にアクセント核がない場合のみ、後続の文節のアクセント核が現れる。すなわち、この様子は表 2.4 のようになる。

### 2.2.6 アクセント結合規則適用則

これまでに述べたアクセント結合規則は、以下のような規則にしたがって適用される。ここで述べる規則は、句坂らによるものであり、GalateaTalk 特有の取り扱いは含まない。

#### 巡回的適用則

単語が複数個連なって文節が作られる場合（例: 食べられませんから = 食べる + られる + ます + ん + から）には、アクセント結合規則は原則的に左から順に適用される。

#### 音節内移動規則

結合規則によって核の移動が生じて 2 モーラ以上からなる音節が核を持つ場合、核はその音節の 1 番目のモーラへ移動する。

#### 無声化に伴う移動規則

無声化した母音が核を持つことになる場合、核は 1 つ左のモーラへ移動する。

#### 一段活用動詞処理規則

動詞のアクセント型の基本を終止形によって決める方法と連用形によって決める方法があるが、終止形を基本とする場合、一段活用（上一段・下一段）の未然形・連用形は終止形よ

<sup>2</sup>複数の文節の結合については、文節内のアクセント結合とは質が異なり、一方のアクセント核の表現が弱まるだけだとする意見もある。

りモーラ数が1つ少ないためアクセント核も左へ1つ移動する(例: シラベル+タ シラベ  
タ)。連用形を基本とする場合はこのような規則は不要で、終止形の活用語尾「る」を様式  
F3で結合アクセント価0の付属語として考えればよい。

#### 先行品詞による結合アクセント価選択則

付属語のうち、様式F2のものの中には、結合アクセント価が先行する自立語の品詞に  
よって異なるものがある。

## 2.3 音声合成システムにおける句坂らの規則の実装

### 2.3.1 GalateaTalkの構成

第1章で述べたGalateaTalkは、擬人化音声対話エージェントツールキット“Galatea  
Toolkit”で使用されている音声合成モジュールである。そのため、顔画像との同期して発  
声する機能を持つなど音声対話を目的とした仕様になっているが、単体でも日本語テキ  
スト音声合成システムとして動作することが可能である。プレーンテキストを入力して読み  
上げさせることが可能だが、それに加え、話者や音量などに関するタグを付けたテキスト  
(JEIDA-62-2000 日本語テキスト音声合成用記号の規格に基づいたもの)にも対応している。  
ソースコードが公開されており無償で使用できる。2007年1月末現在で使用可能な最新バ  
ージョンは2006年10月版である。

GalateaTalkは次の各部分から構成される。

茶筌：形態素解析

ChaOne：複合語処理, 音韻交替処理, アクセント処理

Gtalk：コマンド解析, 韻律・波形生成, 音声出力

このうち“茶筌”は奈良先端科学技術大学院大学にて開発された形態素解析システムであり、  
やはりソースが公開されていて無償で使用できる。

茶筌単体で公開されているものは辞書としてIPA品詞体系辞書(IPADIC)が使用されて  
いるがこれはアクセント情報を持っておらず使用できないため、GalateaTalkではアクセ  
ント情報を持つ独自の辞書“UniDic”を用意し、茶筌にこれを参照させている。

### 2.3.2 Gtalk, UniDicの品詞体系

UniDicは品詞・活用種類・活用形の体系として、およそいわゆる学校文法に合わせた体  
系を使用しているが、一部、学校文法と異なる点がある。特筆すべき点を以下に述べる。

#### 品詞・活用種類・活用形の細分化

単に名詞, 動詞, 形容詞,...などと分けるのではなく、“名詞-普通名詞-サ変可能”, “動詞-一  
般”などのように細分化して分類されている。活用種類(UniDic内では活用型(かつようが



た)と表現される)についても“五段-力行-一般”などとされ、同様に活用形(かつようけい)は“連用形-促音便”などとされる。

#### 形状詞など

学校文法でいういわゆる形容動詞(「変だ」「綺麗だ」など)は、語幹部分を形状詞(“形状詞-一般”など)として扱い、「だ」の部分は助動詞とされる。また、語幹部分が名詞としても使用される場合は“名詞-形状詞可能”とされる。

また、サ変動詞「勉強する」などでは「勉強」を“名詞-サ変可能”として、「する」を“動詞-非自立可能”として扱う。

#### 意志推量形

「書こう」などは学校文法では「書こ」(未然形) + 「う」として扱われるが、UniDicでは全体で“意志推量形-一般”などとする。意志推量形を活用形のひとつとしたのは、話し言葉では「書こ」「書こっ」となることもありこれらの形にも対応する必要があるため、とされる。

### 2.3.3 Gtalk, UniDic における先行品詞別属性選択

2.2 節では、様式 F2 の付属語の一部に、先行する自立語の品詞によって結合アクセント価が異なるものがあるという点について述べた。しかし、UniDic ではすべての付属語に対し、前の語が名詞である場合・動詞である場合・形容詞である場合の 3 つのアクセント属性(アクセント様式と結合アクセント価)を独立して持たせている。

## 2.4 田中による付属語アクセント

比較のため、田中による付属語アクセントの研究 [17] について述べる。

自立語に付属語が続く様子を分析し、句坂らの付属語アクセント結合規則とは異なる方法で付属語を分類したものである。田中の研究は工学的な処理を目的とした句坂らの研究とは異なり、純粋に言語学の立場からのものであり、コンピュータシステムとしての実装はなされていない。

田中の研究によれば、付属語のアクセントは「前の語からの続き方」と「付属語内部のアクセント」の 2 つの要素を持つという。そして、付属語を、「前の語からの続き方」によって表 2.5 に示すように 5 種類の「式」に分け、これらの「式」は以下のような意味を持つものとしている。

従接式 前の自立語にそのまま続くもの。前の語が起伏型の場合には付属語内のさがりめはほとんど目立たなくなる。ただし、前の語が付属語でそこまでに下降が無い場合には下がって続く。

声調式 前の自立語(動詞・形容詞)のトーンが及ぶもの。前接自立語が平板式なら平板式、起伏式なら起伏式になる。

表 2.5: 田中による付属語アクセントの「式」

	従接式	声調式	独立式	下接式	支配式	共下式
名詞接続	/='wa/ /='ga/ /='no/ /='kara/ /='ma <sup>1</sup> de/ ...		/#Qte(te)/		/×gu <sup>1</sup> rai/ /×rasi <sup>1</sup> i/	
動詞接続		/→se <sup>1</sup> ru (sase <sup>1</sup> ru)/ /→re <sup>1</sup> ru (rare <sup>1</sup> ru)/ /→i <sup>1</sup> nai/ /→ta <sup>1</sup> i/ /→na <sup>1</sup> qara/ ...	/#Qte(te)/ /#ta(da)/ /#cja'u (dja'u)/ /#ni/(目的) /#na/(念押) /#so <sup>1</sup> oda/ (伝聞)...	/∟na/ (禁止) /∟'wa/ (取立) /∟mo/ /∟ba/ /∟kara/ ...	/×ma <sup>1</sup> su/ /×'o('jo <sup>1</sup> o)/ /×ta <sup>1</sup> kute/ /×nasa <sup>1</sup> i/ /×rasi <sup>1</sup> i/	
形容詞接続		/→so <sup>1</sup> oda/ (様態)	/#Qte/ /#ta/ /#ba/ /#daro <sup>1</sup> o/ /#mi <sup>1</sup> tai/ /#so <sup>1</sup> oda/ (伝聞)...	/∟te/ /∟'wa/ (取立) /∟mo/ /∟kara/ /∟kedo/ ...	/×rasi <sup>1</sup> i/	

※「斜線封鎖欄」はそこに入る語の可能性がないことを表す。

独立式 前の語からアクセント上独立するもの。従接式と同様だが、付属語には下がらずにつく。

下接式 前の自立語が平板型なら下がって続き、起伏型なら下がらずに続くもの。

支配式 前の語のアクセントに関わりなく、自身の型に引きつけてしまうもの。

なお、田中は、動詞や形容詞のアクセントについては、アクセント核の位置ごとに分類すべきものではなく、下降のない「平板式」と下降のある「起伏式」の2種のトーンにのみ分類するべきであるとの立場に立っているため、付属語の「式」の定義にもそれを前提とした説明がなされているので、注意が必要である。「式」はおよそ句坂らのアクセント結合様式に対応し、「付属語内部のアクセント」は結合アクセント価に対応するものと言える。

## 2.5 長野らによる確率モデルを用いた読み及びアクセント推定

長野ら [20] によって、*n*-gram モデルを用いて、を形態素解析と同時にアクセントを求める研究が行なわれている。

日本語音声合成システムで一般的な手法では、形態素解析によって文を形態素に分割して品詞と読みを決定した後に、アクセント句決定ルールを用いて文を1つ以上のアクセント句の列にし、それぞれのアクセント句について各形態素の単独でのアクセントとアクセント結合規則によって句のアクセントを求める。

表 2.6: 複合名詞「京都タワー」における「京都」及び「タワー」のアクセント

表層	$w$	京都	タワー
品詞	$t$	固有名詞	一般名詞
読み	$s$	kyo : to	ta wa :
アクセント	$a$	L H H	H L L

表 2.7: 複合名詞「京都タワーホテル」における「京都」「タワー」及び「ホテル」のアクセント

表層	$w$	京都	タワー	ホテル
品詞	$t$	固有名詞	一般名詞	一般名詞
読み	$s$	kyo : to	ta wa :	ho te ru
アクセント	$a$	L H H	H H H	H L L

表 2.8: 長野らによるアクセント処理の精度

	単語境界&読み $\langle w, s \rangle$	単語境界&読み&アクセント $\langle w, s, a \rangle$	アクセント $\langle w, s, a \rangle   \langle w, s \rangle$
4つ組による $n$ -gram	97.45%	90.26%	92.63%
規則	97.69%	89.53%	91.65%

それに対し、長野らの手法では、表 2.6 や表 2.7 に示すように、表層（仮名漢字表記） $w$ 、品詞  $t$ 、読み  $s$ 、アクセント  $a$  の 4 つ組を 1 つの単位  $u$  とする。このアクセントは、形態素を単独で発声する場合のアクセントではなく、この 4 つ組  $n$ -gram を  $M_u$  と表すものとする。このモデルにおける 4 つ組単位  $u$  の列の生成確率  $M_u$  は次の式のようになる。

$$M_u(u_1 u_2 \cdots u_h) = \prod_{i=1}^{h+1} P(u_i | u_{i-1} \cdots u_{i-2} u_{i-1})$$

学習データを基に、最適な  $P$  の組み合わせを決定する。そして、入力されたテストデータに対し、形態素列の表層を結合させた文字列と基の文が一致するとの制約条件の下で、 $M_u$  を最も高くする  $u$  の列を探索することで、形態素解析とアクセントが決定される。

この手法の評価を行なうため、長野らは、実際の学習・推定および規則に基づいたアクセント処理との比較を行なった。学習・推定に使用されたコーパスは、新聞記事・テレビニュースの書き起こし・電話応答文など雑多な内容を含み、書き言葉のみならず話し言葉も含むコーパスであるとされる。学習に 8,800 文（190,318 形態素）、推定に 150 文（2,130 形態素）を用いた処理の精度が表 2.8 である。

単語境界・読み・アクセントを同時に推定する手法であるため、アクセント処理の精度は単語境界と読みとアクセントがすべて一致したものの数を、単語境界と読みが一致したものの数で割ったもので表されている。

形態素解析処理に相当する処理とアクセント処理を同時に行なえる点で特に優れた研究であるが、4 つ組による  $n$ -gram を使用して学習・推定したアクセント処理精度が、規則によ

る処理精度を上回っている点も評価する上での1つの重要な点である。ただし、学習と推定に使用したコーパスが公開されておらず、どのような人物によってどのような基準でラベリングされたものであるかが不明である。そのため、長野らによる研究を追試することは難しく、実際の処理性能を検証することは不可能である。

## 第3章 アクセント結合規則の改良

### 3.1 基本となる考え方

ここでは、句坂らによるアクセント処理規則の改良について考える。アクセント結合によるアクセント変化について、その意味を考えることによって系統的で無駄がなくより理解しやすい規則を作ることができるだろうと考え、規則の改良に向けた検討をおこなった。

#### 3.1.1 句坂らの「様式」と田中の「式」の比較

まず、前章において述べた句坂らの付属語結合規則における「アクセント結合様式」と田中による付属語アクセントの「式」の比較をおこなう。田中が付属語の「さがりめ」としてものを句坂らの「結合アクセント価」と同一視した上で両者を対応付けると、表 3.1 のようになる。

句坂らの規則では付属語の連鎖が考慮されていないため、F1, F2 が従接式と独立式のどちらに対応しているとも捉えられるが、いずれも完全な対応ではない。

#### 3.1.2 様式間の共通性への着目

第2章では、単語や文節の結合をおこなった結果としてアクセント型がいかんして決定されるかを、結合の内容別に示した。例えば、自立語と付属語の結合においては後続の付属語によってアクセント結合に関する性質が異なり、それをアクセント様式と称して場合分けしたのだった(表 2.1)。

これは実に有効な場合分けであるが、いささか分類が機械的であるため、少し意味づけを加えてみる。

様式 F1 では、先行する自立語のアクセント型がそのまま文節のアクセント型になるとされる。これについて自立語に核がある場合とない場合に分けて考えてみると、核がある場合

表 3.1: 句坂らの「様式」と田中の「式」の比較

田中の「式」	句坂らの「様式」
従接式	F1, F2@{0 以外 }
独立式	
下接式	F2@0
声調式	F3@{X} <sup>1</sup>
支配式	F4@{0 以外 }, F5

<sup>1</sup>声調式は起伏式の語(下降のある語)に結合すると全体として起伏式になるということを示すのみであり、様式 F3 の特定の結合アクセント価とは対応付けられない。

表 3.2: 書き換えた付属語結合規則

様式	$M_1$	$\widetilde{M}_2$	$M_c$
F1	$\phi$ $x$	$[\phi]$ —	$\phi$ $x$
F2	$\phi$ $x$	$y$ —	$N_1 + y$ $x$
F3	$\phi$ $x$	$[\phi]$ $y$	$\phi$ $N_1 + y$
F4	$\phi$ $x$	$y$ $y$	$N_1 + y$ $N_1 + y$
F5	$\phi$ $x$	$[\phi]$ $[\phi]$	$\phi$ $\phi$
F6	$\phi$ $x$	$y_a$ $y_b$	$N_1 + y_a$ $N_1 + y_b$

表 3.3: 単純化した付属語結合規則

様式	$M_1$	$M_c$
F1,2	$\phi$ $x$	$N_1 + \widetilde{M}_2$ $x$
F4,5	$\phi$ $x$	$N_1 + \widetilde{M}_2$ $N_1 + \widetilde{M}_2$
F3,6	$\phi$ $x$	$N_1 + \widetilde{M}_{2a}$ $N_1 + \widetilde{M}_{2b}$

は付属語を発音し始める段階ですでに低になっているが、核がない場合はその段階で高のままである。つまり、自立語に核がない場合は付属語の中で低に移る変化の可能性が残されているにもかかわらず、それでも最後まで低に移らないのがこの様式である。すなわち、F1においてもF2同様に、「先行する自立語に核がある場合はそちらが残り、自立語に核がない場合は付属語の持つアクセントによって核が決定される。」といえ、F1の場合は結合アクセント価が核なしを表す  $\phi$  価として考えればよいのである<sup>2</sup>。

これに基づいて表 2.1 を書き換えると、表 3.2 のようになる。 $x$  と  $y$  は  $\phi$  でない有限の値を表す。ただし、 $\widetilde{M}_2$  については、もともとの規則では定義されないところについても、適切だと思われるものを括弧に入れて補完してある。

そして、 $\phi$  について、計算式で記述する場合にも都合がいいように演算を「 $n + \phi \rightarrow \phi$ 」と形式的に決めることにすると、表 3.2 は表 3.3 のように 3 つの様式にまとめることができる。様式 F1 と F2、F5 と F4、F3 と F6 はそれぞれ本来の“様式”として持つ性質は共通であり、単に結合アクセント価が異なるだけ（ $\phi$  価であるかないか）といえる。

これがただちに規則の精度の向上に結びつくわけではないが、以降で活用による結合アクセント価の変化について考える際に、このそれぞれのペアを共通の様式ととらえられる点を用いる。

### 3.1.3 活用の考慮

動詞・形容詞などの用言や助動詞では、活用によって単語の形が変化する。しかし、句坂らの規則や Gtalk, UniDic で用いている規則では、活用がアクセントに与える影響についてはほとんど触れておらず、唯一、一段活用動詞処理規則のみがこれに該当する。

<sup>2</sup>本論文ではアクセント核を持たない型（平板型）をあえて一般的でない  $\phi$  型として記述してきたが、その理由はここにある。一般には平板型を 0 型と表すが、0 型という表現は本来 0 の位置に核がある様子（ $\overset{\cdot}{\quad}$ ；結合アクセント価 0 といったときはこれを表している）を表すのに都合が良いものであるため、適切とはいえない。  $\phi$  を用いれば結合アクセント価について 0 価と  $\phi$  価の双方が表現できることも大きな利点である。なお、同様の理由により、 $\emptyset$  や  $\infty$  が用いられることがある [13, 14]。

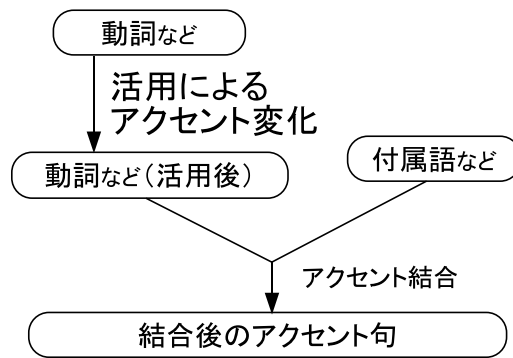


図 3.1: 活用によるアクセント変化の考慮

活用によるアクセント変化があると考え、アクセント結合の処理をする際には図 3.1 のようにまず先行する用言・助動詞の活用による変化の後のアクセント型を求め、それを先行語の型として後続の語との接続をして結合後のアクセントを求めることになる。

### 3.1.4 先行品詞別属性選択の是非

第 2 章で述べたように、Gtalk, UniDic では付属語のアクセント属性を先行語が名詞・動詞・形容詞のそれぞれの場合で分けている（これを以降では名詞用・動詞用・形容詞用と呼ぶことにする）が、このように区別することが本当に必要であるのか、また、このような区別だけでは不十分でないのか、という疑問がある。例えば、先行語が助動詞の場合の処理が規定されていない点は問題である。

動詞・形容詞・助動詞の 3 品詞は、活用をする点や後続する語の種類などに共通性が見られるため、これらに対して同一のアクセント属性が使えるのではないかと予想が可能である。UniDic において先行語が動詞の場合と形容詞の場合で異なる値を持つことがあったのは、活用によるアクセント変化を考慮していないことなどによるものではないかと予想できる。

## 3.2 活用による影響の検討

### 3.2.1 動詞の活用

動詞の活用がアクセントに及ぼす影響を検討するため、動詞に付属語等を結合させた後のアクセントを調べ、そこから、その付属語等を結合させた場合のアクセント型として正しい値が算出されるような動詞部分のアクセント型を逆算する推定作業を、各活用種類ごとに具体的な単語を少数用いておこなった<sup>3</sup>。推定のために用いる付属語等およびそのアクセント様式として、

<sup>3</sup>筆者は幼児期から現在に至るまで東京在住であり、東京アクセント話者である。

表 3.4: 動詞-五段活用 (歩く)

	結合後	活用変化
未然	アルカ・ナイ	アルカ?
連用	アルイ・タ	アルイ
終止	アルク	アルク
連体	アルク・コト	アルク
仮定	アルケ・バ	アルケ
命令	アルケ	アルケ

表 3.5: 動詞-五段活用 (笑う)

	結合後	活用変化
未然	ワラワ・ナイ	ワラワ
連用	ワラッ・タ	ワラッ
終止	ワラウ	ワラウ
連体	ワラウ・コト	ワラウ
仮定	ワラエ・バ	ワラエ
命令	ワラエ	ワラエ

表 3.6: 動詞-一段活用 (見る)

	結合後	活用変化
未然	ミ・ナイ	ミ
連用	ミ・タ	ミ
終止	ミル	ミル
連体	ミル・コト	ミル
仮定	ミレ・バ	ミレ
命令	ミロ, ミヨ	ミ+“ロ, ヨ”?

表 3.7: 動詞-一段活用 (生きる)

	結合後	活用変化
未然	イキ・ナイ	イキ?
連用	イキ・タ	イキ
終止	イキル	イキル
連体	イキル・コト	イキル
仮定	イキレ・バ	イキレ
命令	イキロ, イキヨ	イキ?+“ロ, ヨ”?

表 3.8: 動詞-一段活用 (居る)

	結合後	活用変化
未然	イ・ナイ	イ
連用	イ・タ	イ
終止	イル	イル
連体	イル・コト	イル
仮定	イレ・バ	イレ
命令	イロ	イ+“ロ”?

表 3.9: 動詞-一段活用 (借りる)

	結合後	活用変化
未然	カリ・ナイ	カリ
連用	カリ・タ	カリ
終止	カリル	カリル
連体	カリル・コト	カリル
仮定	カリレ・バ	カリレ
命令	カリロ, カリヨ	カリ+“ロ, ヨ”?

未然形：ない (様式 F3,0 価), れる (様式 F3,1 価), ず (様式 F3,0 価)

連用形：た (様式 F2,1 価)

連体形：こと (文節)

仮定形：ば (様式 F2,0 価)

を用いた。なお、活用形については学校文法に基づいた体系とする。

まず、五段活用においては、表 3.4, 表 3.5 のようになった。表の左の列は付属語等との結合後のアクセント、右の列はそこから推定した動詞の活用後・結合前のアクセントである。“?”がある部分は、一意に決定できず他からの類推して選んだものである。終止形で有核のもの・無核のものとも、活用の影響はないと言っていだろう。

一段活用 (上一段・下一段) では、表 3.6 ~ 表 3.9 のようになった。動詞としては上一段活用の動詞で、有核のもの・無核のものそれぞれで 2 モーラのもの・3 モーラのもの、計 4 語を用いた。表中で、命令形の右の列には「カリ+“ロ, ヨ”?’のように記してあるが、これ



表 3.10: 動詞-カ変活用(来る)

	結合後	活用変化
未然	コ・ナイ	コ
連用	キ・タ	キ
終止	クル	クル
連体	クル・コト	クル
仮定	クレ・バ	クレ
命令	コイ	コイ

表 3.11: 動詞-サ変活用(する(追求する,...))

	結合後	活用変化
未然	シ・ナイ サ・レル セ・ズ	シ サ セ
連用	シ・タ	シ
終止	スル	スル
連体	スル・コト	スル
仮定	スレ・バ	スレ
命令	シロ セヨ	シ + “ロ”? セ + “ヨ”?

表 3.12: 動詞-サ変活用(発する(感ずる,...)) 表 3.13: 動詞-サ変活用(利する(信ずる,...))

	結合後	活用変化
未然	ハッシ・ナイ ハッセ・ズ	ハッシ, ハッセ
連用	ハッシ・タ	ハッシ
終止	ハッスル	ハッスル
連体	ハッスル・コト	ハッスル
仮定	ハッスレ・バ	ハッスレ
命令	ハッシロ ハッセヨ	ハッシ + “ロ”? ハッセ + “ヨ”?

	結合後	活用変化
未然	リシ・ナイ リゼ・ズ	リシ?, リゼ?
連用	リシ・タ	リシ
終止	リスル	リスル
連体	リスル・コト	リスル
仮定	リスレ・バ	リスレ
命令	リシロ, リゼヨ	リシ? + “ロ”? リゼ? + “ヨ”?

はそのように分離して付属語“ロ”(様式 F3,0 価)や“ヨ”(様式 F4,0 価)を考えると他の活用形と動詞部分が未然形・連用形と同じようになることを示したものである。この分離をしないのであれば単純に左の列の形が命令形のアクセントとなる。

ここから、2モーラの語では変化は無いが、3モーラ(以上)の語では有核の動詞の連用形で核の移動があることが分かる。これは第2章で述べた一段活用動詞処理規則である。また、“ロ、ヨ”の分離をしないのであれば、無核の動詞の命令形のうち「カリヨ」のようなものでは核が発生することといえる。

続いてカ行変格活用だが、これは表 3.10 のようになった。変化は無いといって良いだろう。

そして、サ行変格活用は表 3.11~表 3.13 のようになった。単純な「する」の他に、「単漢字+する」で無核のもの、同じく有核のもの3種類について調べた。単純な「する」の場合のみに「セ・ズ」「セ・ヨ」で核が生じ、また、「単漢字+する」では有核のもの・無核のものとも一段活用と同様の傾向を示した。

上記のように、一段活用動詞処理規則だけでは活用の影響に関する処理は不十分であり、一段活用命令形およびサ変活用の一部にも対処する必要があるということが判明した。

### 3.2.2 形容詞の活用

次は、形容詞についてである。後続させる付属語等については、動詞のときと同様のものを用い、結合アクセント価も同値と仮定する。

さて、形容詞には有核のものと無核のものがあり<sup>4</sup>、有核のもの核は基本的に語幹の最終モーラ（活用語尾“い”の手前）である。「オ<sup>1</sup>オ<sup>1</sup>イ」は音節内移動規則により「オ<sup>1</sup>オ<sup>1</sup>イ」から核が移動したものと見えるため、そのように考えればこれも同様に扱える。

そのそれぞれについて3モーラの語と4モーラの語について調べた結果が表3.14～表3.17である。なお、未然形については様式F4の「う」以外は後続しえないため、形容詞単独の型は判断できない。

まず、有核のものについて見ると、3モーラのものでは連用形と仮定形で核が前に移動する。ただし、移動させなくても不自然でなく、重要度は低い。

無核のものでは、連用形の一部と仮定形において語幹の最終モーラに核が発生する。ただし、ものによっては核を発生させなくても許容でき、終止形でも核を発生させた形に不自然さは無いなど、核の存在の可否は極めて不安定である。このように無核の形容詞に核が発生することは、付属語の結合アクセント価を「前の語が形容詞の場合には-2価」などとすれば対応可能であり、実際にUniDicでは付属語にそのような値を持たせている。しかし、ここではむしろ、核の発生が形容詞の持つ性質であると捉えれば、形容詞にも動詞用の結合アクセント価がおよそ使用可能である点に注目すべきである。

このように、形容詞についてははっきりとした正解の示し難いものではあるが、いずれにせよ核の移動・発生を考慮する必要があることが分かった。また、形容詞の活用に関しては、3モーラの有核形容詞についてのものと無核形容詞についてのものの2種類に大きく分かれることに注意する必要があるだろう。

### 3.2.3 助動詞の活用

自立語である動詞・形容詞の他に、付属語の助動詞も活用をする。助動詞の多くは動詞や形容詞と同様の活用をするため、動詞・形容詞について述べたことがそれぞれ動詞型活用・形容詞型活用の助動詞にも当てはまるはずである。したがって、これまで述べたようなアクセント変化や、後続語の結合アクセント価として動詞用の値が使える点である点が、助動詞についても言えるだろう。また、自立語に関しては活用によるアクセント型の変化を見たが、助動詞の場合は同じように活用で結合アクセント価が変化すると考えて良いだろう。

しかし、活用によるアクセント変化を考える際、アクセント型の $\phi$ 型に相当する結合アクセント価が第3.1節で述べた $\phi$ 価であるということに注意する必要がある。すなわち、句坂らの規則で様式F1とされる助動詞は、“様式F2, $\phi$ 価”であるものとして考えることとなる。様式F5とF4、F3とF6についても同様である。なお、結合アクセント価の0価に相当するアクセント型が無いが、動詞や形容詞と同様の活用をする0価の助動詞はUniDic中では「ちゅう」「という」の変化したもの）等しか見当たらず、変化を示さないと考えて良いだろう。

それ以外の特殊な活用の助動詞について、個別に見る必要がある。断定の助動詞「だ」について考えたものが表3.18である。後続する付属語の結合アクセント価は、やはり動詞に

<sup>4</sup>終止形単独では分かりづらいが、後に「こと」「もの」「ひと」などが続くときと違いがはっきりする。

表 3.14: 形容詞-現代形容詞活用 (強い)

	結合後	活用変化
未然	ツヨカ <sup>レ</sup> 口・ウ	?
連用	ツヨカ <sup>ッ</sup> ・タ ツヨク・ナル ツヨク・テ ツヨ <sup>ー</sup> ・ゴザル	ツヨカ <sup>ッ</sup> ツヨク ツヨク ツヨ <sup>ー</sup>
終止	ツヨ <sup>イ</sup>	ツヨ <sup>イ</sup>
連体	ツヨ <sup>イ</sup> ・コト	ツヨ <sup>イ</sup>
仮定	ツヨケ <sup>レ</sup> ・バ	ツヨケ <sup>レ</sup>

表 3.15: 形容詞-現代形容詞活用 (鋭い)

	結合後	活用変化
未然	スルドカ <sup>レ</sup> 口・ウ	?
連用	スルドカ <sup>ッ</sup> ・タ スルドク・ナル スルドク・テ スルド <sup>ー</sup> ・ゴザル	スルドカ <sup>ッ</sup> スルドク スルドク スルド <sup>ー</sup>
終止	スルド <sup>イ</sup>	スルド <sup>イ</sup>
連体	スルド <sup>イ</sup> ・コト	スルド <sup>イ</sup>
仮定	スルドケ <sup>レ</sup> ・バ	スルドケ <sup>レ</sup>

表 3.16: 形容詞-現代形容詞活用 (赤い)

	結合後	活用変化
未然	アカカ <sup>レ</sup> 口・ウ	?
連用	アカカ <sup>ッ</sup> ・タ アカク・ナル アカク・テ アコ <sup>ー</sup> ・ゴザル	アカカ <sup>ッ</sup> アカク アカク アコ <sup>ー</sup>
終止	アカ <sup>イ</sup> (アカ <sup>イ</sup> )	アカ <sup>イ</sup> (アカ <sup>イ</sup> )
連体	アカ <sup>イ</sup> ・コト	アカ <sup>イ</sup>
仮定	アカケ <sup>レ</sup> ・バ	アカケ <sup>レ</sup>

表 3.17: 形容詞-現代形容詞活用 (危ない)

	結合後	活用変化
未然	アブナカ <sup>レ</sup> 口・ウ	?
連用	アブナカ <sup>ッ</sup> ・タ アブナク・ナル アブナク・テ アブ <sup>ノ</sup> ・ゴザル	アブナカ <sup>ッ</sup> アブナク アブナク アブ <sup>ノ</sup>
終止	アブナ <sup>イ</sup> (アブナ <sup>イ</sup> )	アブナ <sup>イ</sup> (アブナ <sup>イ</sup> )
連体	アブナ <sup>イ</sup> ・コト	アブナ <sup>イ</sup>
仮定	アブナケ <sup>レ</sup> ・バ	アブナケ <sup>レ</sup>

表 3.18: 助動詞-ダ活用 (だ)

	結合後	活用変化
未然	…ダ <sup>レ</sup> 口・ウ	?
連用	…ダ <sup>ッ</sup> ・タ …デ、 …デ・モ (デと同じ)	…ダ <sup>ッ</sup> …デ? … … … <sup>ニ</sup>
終止	…ダ。 …ダ・カラ	…ダ
連体	…ナ・コト	…ナ
仮定	…ナ <sup>ラ</sup> 、 …ナ <sup>ラ</sup> ・バ	…ナ <sup>ラ</sup>

表 3.19: 助動詞-デス活用 (です)

	結合後	活用変化
未然	…デ <sup>シ</sup> 口・ウ	?
連用	…デ <sup>シ</sup> ・タ	…デ <sup>シ</sup>
終止	…デ <sup>ス</sup>	…デ <sup>ス</sup>
連体	…デ <sup>ス</sup> ・コト	…デ <sup>ス</sup>
仮定	…デス <sup>レ</sup> ・バ	…デス <sup>レ</sup>

後続する場合と同値であるとした。ただし、接続上の都合から後続語として「も」「さえ」「から」「と」(様式 F2,0 価)を新たに用いた。様式 F2,0 価の付属語を後続させたもので後続語の直前が核になったものでは、「だ」の部分の有核か無核か判断しづらいが、明らかに核の発生したところとして連用形の「ダッ・タ」や仮定形の「ナラ・バ」があり、これを捉えておけば問題ないだろう。なお、過去の助動詞「た」も類似した形をとる。

また、丁寧の助動詞「です」では表 3.19 のように仮定形で核の移動が発生する。丁寧の助動詞「ます」についても同様である。

### 3.3 活用を考慮したアクセント推定実験

このように活用によるアクセント変化を考慮することで、アクセント推定の正解率をどの程度向上させるか、および逆に新たな問題が発生することがないかを検証するため、活用によるアクセント変化を処理するモジュールを GalateaTalk に組み込んだものを用意し、これと元の GalateaTalk に、毎日新聞朝刊記事の 1995 年 1 月 1 日から 9 日分を文節単位に分けたもの(有効な文節数: 85,413 文節)を入力して、文節ごとにアクセント型を出力させた。そして、適用する規則の内容の違いにより出力が異なったものについて、個別に変化の内容を調べた。この実験の結果、以下のことが分かった。

- 動詞・3 モーラ有核形容詞の活用及び意志推量形に関する核の移動・発生 の考慮では、正答率が計 0.5% 程度向上し、重大な悪化は起きない。
- 無核形容詞の核の発生については、同じ活用形でも「アカク」「アカク・テ」と異なるため、活用形ごとの判断では対応しきれない。この連用形「...ク」は、核を発生させない方が出現頻度の点から正答率が高くなる。活用形ごとでなく「付属語が後続するときは核が発生する」とする方法で核の発生を判別する方法を用いたほうが正答率が高くなるが、依然「アカイ・クライ」等、対応できない例がある。活用形ごとに判別し、連用形に関してのみこの方法を用いると最も良い規則となる。
- 助動詞の活用の考慮では、正答率が 0.5% ほど向上した。形容詞型活用のもものでは、前述の形容詞の場合と同様の問題がある。
- GalateaTalk で付属語の先行品詞として動詞・形容詞を区別しているものはほとんど意味を成しておらず、動詞用の属性に統一できそうである。先行語が助動詞の場合にも動詞用の属性を用いて構わない。

この結果、入力文節数の 1.0% 程度において不正解から正解に変化しており、重大な悪化は起きないことが分かった。

### 3.4 その他の課題点

#### 副次アクセントを扱えない

アクセント句に複数のアクセント核(副次アクセント)が存在する場合のあることが指摘されているが[7, 13]、現在の GalateaTalk では 1 つのアクセント句には高々 1 つのアクセント核を持つことしか許されず、副次アクセントが表現できない。例えば、

- アヤシ 'カッタヨ 'ーダ (怪しかったようだ)
- タノ 'ンデミ 'テクダサイマセ 'ンデシヨ 'ーカ (頼んでみてくださいませんか)

(“ ’ ” の位置がアクセント核) のような場合に副次アクセントが発生するとされる。

この点は、アクセント句の概念がまだ明確にされていないことが原因とも言えるが、明確に定義づけるのは難しいものと思われる。

### 規則そのものや規則の実装に不十分な点がある

句坂らの規則を基にしている現在の GalateaTalk には不十分な点があり、例えば、田中によって示された付属語アクセント [17] では付属語の連鎖(「私には」など)が考慮されているのに対し、句坂らの規則ではそれに対応できないなどの欠点が見られる。また、句坂らの規則のうち接頭辞の規則については、2 語の結合の結果複数のアクセント核が残るとされるものがあるが、アクセント結合の考え方と相容れないものであり、前述のとおり GalateaTalk の実装ではアクセント句内に複数の核を持つことが許されないために対応できていない。

### 例外的なアクセント処理を記述できない

現在の実装では、「大きさ」が“  $\overline{\text{オーキサ}}$  ”ではなく“  $\overline{\text{オーキサ}}$  ”が正解となる、「今まで」が“  $\overline{\text{イママデ}}$  ”ではなく“  $\overline{\text{イママデ}}$  ”が正解となるなど、例外的なアクセントを示すものについての処理を記述することが容易ではない。

GalateaTalk で用いている辞書 UniDic では、可能な限り細かく分けたものとして単語を捉えるため、少ない規則によるアクセント結合処理では誤りが多くなりやすく、規則からの例外となるパターンも多く存在すると考えられる。

## 第4章 単独の日本語話者によるアクセントデータベースの作成

### 4.1 アクセントデータベース作成の目的

これまでに述べたように、規則に基づいてすべての事象を網羅することには限界があり、これを改善する方法の1つとして統計的な学習・推定が考えられる。統計的手法を導入すれば、多くの事象について幅広く取り扱えるようになることが期待でき、例外的なものも含めた多くのアクセント処理の実現が期待できる。しかし、統計的な手法で必要となる自由な利用が可能な多量の正解ラベル付きデータベースは存在しておらず、難しい状況にある。また、同じ理由で、規則による処理がどの程度妥当であるかも判定することができない問題もある。

以上のようなアクセント研究上の問題を解決するために、アクセントデータベースの作成を計画し、実行している。本論文の執筆時点では全体のおよそ四分の一が完成しており、以後も継続的に続ける予定である。

### 4.2 アクセントデータベースの仕様

#### 4.2.1 ラベリング内容

1人のラベラによって、

- 文発声時についての“アクセント句境界位置”と“アクセント核位置”
- 文中に現れる自立語の単独発声時についての“アクセント核位置”

についてラベリングしたを多量に作成し、データベースとすることとした。詳細については後述する。

#### 4.2.2 ラベリング対象とする言語事象の定義

第2章で述べたとおり、日本語東京方言で文を発声する際、文はいくつかのまとまりに分かれ、それぞれの内部では音の高さが連続的に変化する。まとまりが始まる箇所ではピッチ（音の高さ）の上昇が見られ、その後、まとまりの内部では上昇は無く、ゆるやかに下降してゆく。また、まとまりの内部には、語彙に依存した比較的急激な下降箇所がおおむね高々1つ存在する。このようなまとまりをアクセント句と呼び、急激な下降の箇所をアクセント核とするのが一般的である。

しかし、実際の発声におけるピッチ変化を観察すると、明確にアクセント句の分離ができるわけではなく、複数のアクセント句が影響を及ぼし合い、また、融合する現象が見られる。そのため、アクセント句の認定は明確にしづらく、また、各アクセント句の内部には高々1

つのアクセント核しか存在しないとする考え方も危うくなり、複数のアクセント核が存在しても良いと考えねばならない。

アクセント句内に複数のアクセント核が存在できるものとする場合には、音の高さを高・低の2値と捉えることはせず、何段にもわたって下降が起こる現象が存在するものと認識し、その下降箇所をすべてアクセント核として扱う。川上 [13] による例を用いると、2つの文節が一息で発音される場合、少なくとも一方が平板式のアクセントを持つ場合には、

- ニワニ サイタ (平板式+平板式; 庭に咲いた)
- サイタ オミナ 'エシ (平板式+起伏式; 咲いた女郎花)
- オミナ 'エシガ サイタ (起伏式+平板式; 女郎花が咲いた)

のようになる ([ ' ] はアクセント核による下降の位置)。双方が起伏式である場合についても、同様に、

- オミナ 'エシオ イ 'ケタ (平板式+平板式; 女郎花を活けた)

になるものと考えることができる。この点について、句坂らの規則を含むいくつかの説では第2の文節のアクセント核が消失して

- オミナ 'エシオ イケタ

となるものとしている。このどちらが正しいかは難しい問題であるが、

- キノ 'ーモ フラ 'レタ (昨日も降られた)
- キノ 'ーモ フラレタ (昨日も振られた)

の双方を一息で読んだとしても、それらが類似したものとなるとは言え、区別が全く無い同一のものになるとは言いがたく、1つのまとまり(アクセント句)の中に複数のアクセント核が存在することは認めざるを得ないであろう。なお、複数回の下降が可能とする考え方の中には、2回までの下降が可能とするアクセントの三段観や、同様に3回までの下降を許す四段観などというものがあるが、下降の回数を2回や3回に制限する明確な理由は無く、言語的な観点からは無制限とするのが的確と思われる。

日本語話し言葉コーパス [18] においては、複数の文節で構成されたまとまりが、後続文節の句頭上昇が見られない程度にまで融合していれば1つのアクセント句として扱うことを原則としている。しかし、同程度の融合が見られても双方が核を持つ文節である場合は、「アクセント句には1つのアクセント核しか存在し得ない」との考えに基づき、複数のアクセント句として扱っている。しかし、双方が有核であるばあいのみ融合が起こらないとするのにはいささか不合理であると言わざるを得ない。

実際に、複数のアクセント核を持つように見えるアクセント句は、おおむね、遅く読むなどする場合に複数の(小さな)アクセント句に分かれるものが、1つにまとまったものと考えられる。言い換えれば、2つめ以降のアクセント句の先頭にあったピッチの上昇が無くなるとともに2つめ以降のアクセント句内の核におけるピッチの下降もいくらか弱まった結果、大きな1つのまとまりになったものと言うことができる。このように考えれば、アクセント句は高々1つのアクセント核しか持たないものと決めことによって小さなアクセント句を認定し、それらがゆるやかに連結するのだと考えることができる。しかし、その場合には、一息で読んだ

- ニワニ サイタ (平板式+平板式; 庭に咲いた)

はアクセント核を持たないため、いくつかのアクセント句と捉えるべきか判断が難しい。しかし、文節とアクセント句を同一視することでこの問題を解決しようとする、

- ヨ 'ム ダロ 'ー (読むだろう)
- カ 'ンヨ セ 'ヌ (関与せぬ)
- ニ 'ジュー イチ 'バンノ (二十一番の)

のように文節内に複数の下降箇所(アクセント核)を持つものは適切なアクセント句認定とならない。

いずれにせよ、この現象に対応するためには、アクセント句の中に複数の核が存在することを記述できるできるようにするか、逆にアクセント句を小さいサイズ(文節に達しない大きさ)で捉えた上で複数のアクセント句をゆるやかに結合させる仕組みを持たせるかのいずれかが必要となるが、これに対する明確な解答は存在しないのである。

アクセント核の認定についても、アクセント核は発声者や聴取者にとって知覚されているにもかかわらず実際のピッチ変化に明確に現れていない場合が存在し、音の高さの変化だけでは判断できない問題がある[18]。その上、発声者等の側も必ずしも明確にアクセント核の知覚を持っているではない。

このように、アクセント句とアクセント核はいずれも明確に認定できるものでない。特に、1つのアクセント句の中には高々1つしかアクセント核が存在できないものとする場合、アクセント句とアクセント核のどちらを先に認定してからもう一方を認定するかによって結果に大きな影響を与える可能性がある。これらの認定に文法的要素が関係するのであれば、ラベラに対して文法的な観点からの考慮も含めたラベリングをさせることとなるが、その作業はラベラの知覚から逸脱したものとなる可能性があり、不適切である。

そこで、今回のアクセントラベリングにおいては以下のように定義し、これらをラベリング対象とした。

アクセント句 句頭上昇が見られる箇所をアクセント句の境界とする。視覚提示される話速(約7モーラ/秒)に合わせて自然に読んだ場合を想定する。休止が入った場合も、通常はアクセント句の境界が生じる。

アクセント核 音の高さが下がる箇所の直前のモーラ。アクセント句内の出現数は制限しない。意識的に当該箇所で急激に音の高さを下げそれ以外では同じ音の高さで平坦に(イントネーションを除去して)読むようにした場合に、違和感が生じない位置をアクセント核と認める。

この定義により、ラベラは、知識にとらわれず比較的容易に知覚をラベルに直結させることができるものと考えられる。

#### 4.2.3 ラベラを1人とする理由及び単独発声ラベリングの目的

ラベリング作業を行なうのは1人のみとした。文に対するラベリングとしては、文中におけるアクセント句境界位置とアクセント核位置を、1人のみのラベラにラベリングさせるものとした。ラベラを1人のみとしたのは、複数人によるラベリングではアクセント感覚の個



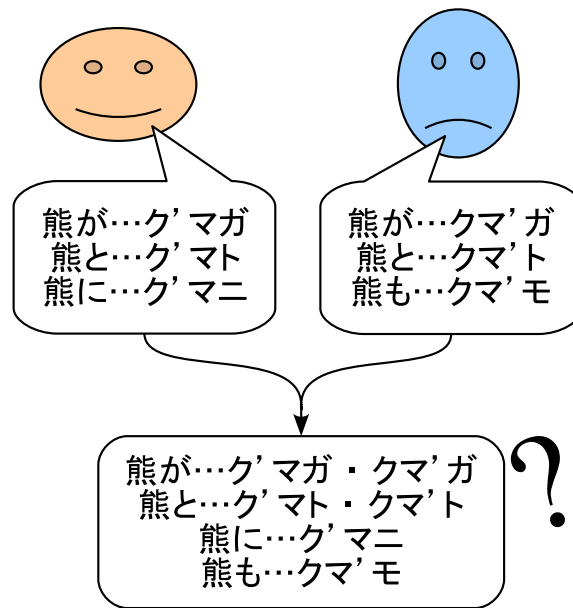


図 4.1: 複数人でラベリングをすることの弊害

人差が入ってしまい、アクセントが変化する現象と個人差によるアクセントの差異が混ざって、アクセント研究用資料としての価値を下げるものとなるからである。図 4.1 に示すように、東京方言話者でも「熊」を「クマ」とする者と「クマ」とする者がおり、それらの混ざったデータは、単純なアクセント現象を極めて複雑なものであるがごとく見せてしまう。

また、文中でのアクセントと個々の形態素を単独で発声する場合のアクセントを比較することによる研究で用いることを想定し、文中に出現するすべての自立語に対して単独発声時のアクセントについてのラベリングを行なわせた。形態素単独発声時のアクセントはアクセント辞典等に記されているが、アクセント感覚の個人差を含まないものとするため、文に対するラベリングを行なったラベラ自身に、改めて単独発声時についてのラベリングをさせた次第である。自立語でない形態素については、単独発声を想定することは難しく、また、不合理であるため行っていない。

#### 4.2.4 使用した文セット

日本音響学会 新聞記事読み上げ音声コーパス (JNAS) [19] で使用されている文 (毎日新聞記事から選ばれた 16,178 文および ATR 音素バランス文 503 文) を用いた。既に音声コーパス用に使用されている文であることから、すべての文に読みが与えられていること、碎けた文が少なく扱いやすいことなどによる。使用されている新聞記事文は、「CD-毎日新聞」の 1991 年版から 1994 年版までの記事から選ばれたものである。

これらの文は、音声読み上げ用に自動的な形態素解析処理と手作業を組み合わせるすべての文に読みが付与されている。ただし、不自然な読みや誤った読みも散見されるため、修正を加えた。修正の方法や基準などについては後述する。

#### 4.2.5 形態素解析

文中に現れる自立語に関するラベリングをおこなうため、また、ラベリング結果を利用する際の利便性のため、すべての文に対する形態素解析結果を用意をした。形態素解析の品詞体系は UniDic[10, 11] に基づくものとした。UniDic は、前述の GalateaTalk で使用されている辞書であり、多くの者に受け入れられる標準的な品詞体系を使用すべきとの観点から、いわゆる学校文法に近い体系が採用されている。テキスト音声合成での利用のためにアクセント型などの値を持っている。アクセントデータベースの作成において UniDic を採用したのは、GalateaTalk に採用されている体系を用いることで、作成したデータベースがアクセント研究において使用しやすいものとなるとの考えに基づくものである。

形態素解析結果は、

- 語彙素（形態素を代表する表記）についての
  - － 漢字表記（“語彙素”，例：「含める」）
  - － 仮名表記（“語彙素読み”，例：「フクメル」）
- 語形（語彙素を語形の違いによって細分化したもの）についての
  - － 仮名表記（“語形”），例：「フクメル」
- 書字形（活用等による変化後の形であり実際に出現した表記）についての
  - － 漢字表記（“書字形”，例：「含め」）
  - － 仮名表記（“仮名形”，例：「フクメ」）

の計 5 種類の漢字・仮名表記に加え、

- 品詞, 例: 「動詞-一般」
- 活用型, 例: 「下一段-マ行」
- 活用形, 例: 「連用形-一般」
- アクセント型, 例: 「3」
- アクセント結合型, 例: 「C1」

の情報を持っている。

この形態素解析作業については、UniDic の開発をしている千葉大学の伝康晴助教授に依頼した。形態素解析作業は、辞書として UniDic を用いて自動的に行った後、用意した読みと異なった読みが付与されたものを中心に手作業での修正が施されている。従って、すべて正確に形態素解析が行われているわけではないが、読みについてはラベリングに用いたものと完全に同一になっている。

#### 4.2.6 読み仮名の修正

読み仮名の修正のため、不自然な読みを指摘する担当のアルバイト学生を用意し、すべての文を確認させた。このアルバイト学生からの指摘および形態素解析依頼先からの指摘などを基に、明らかな誤りは修正し、明らかでないものについても個別に判断して修正を施した。一括して修正したものとして、

- 数字の「2」「二」(ニ)や「5」「五」(ゴ)の後に小数点を表す「・」(テン)または他の数字が続く場合、読み仮名を「ニイ」や「ゴオ」に変更した。逆に、それ以外の場合については「ニ」や「ゴ」で統一した。
- 「横浜マ」などのスポーツチーム名の略記については、「マ」の部分の読みを「マリノス」とするなど、略さない読みを採用した。
- スポーツのポジション名「GK」「FW」などの読みは、「ゴールキーパー」「フォワード」などとした。
- 「米」「仏」などの国名の略記については、漢字を音読みした「ベイ」「フツ」などに統一した。
- 「私」(ワタシ・ワタクシ)など複数の読みが許されるものについては、原則としてJNASでの読みのまま変更をしなかった。

などがあげられる。

#### 4.2.7 ラベリング従事者の選定

ラベリング作業に先立ち、ラベラとラベル検査担当者を選定した。

ラベリング作業をおこなうのは、先に述べた目的に合わせ、1人のみとした。また、多量の文に対しラベリング作業を行なうため誤ったラベルを付与してしまうことは免れ得ないため、ラベリングの済んだすべて文のラベルに目を通し、“誤りが含まれる可能性がある文”を選び出す作業の従事者(1人、今後増員の可能性あり)を、ラベラとは別に用意した。これら2人はいずれも東京生まれ・東京育ちの大学生であり、合唱の団体に所属するため比較的鋭い音感を持っていると推定される。ほぼ同様の学生6人の中から、アクセントに関する簡単な教育をした上で試験によって選抜したものである。

また、これとは別に、前述の読み仮名修正担当者も選定している。

#### 4.2.8 実際のラベリング作業の詳細

ラベラとラベル検査担当者には、付録に掲げた解説文書を熟読させた上で、それに従って作業をするよう求めた。

ラベリング作業は、音声を经ず、テキストに直接ラベルを付与させる形で行なった。音声を経ると極めて多くの作業時間を要するためであり、また、ラベラにとっての知覚的なアクセント情報を重視するためでもある。

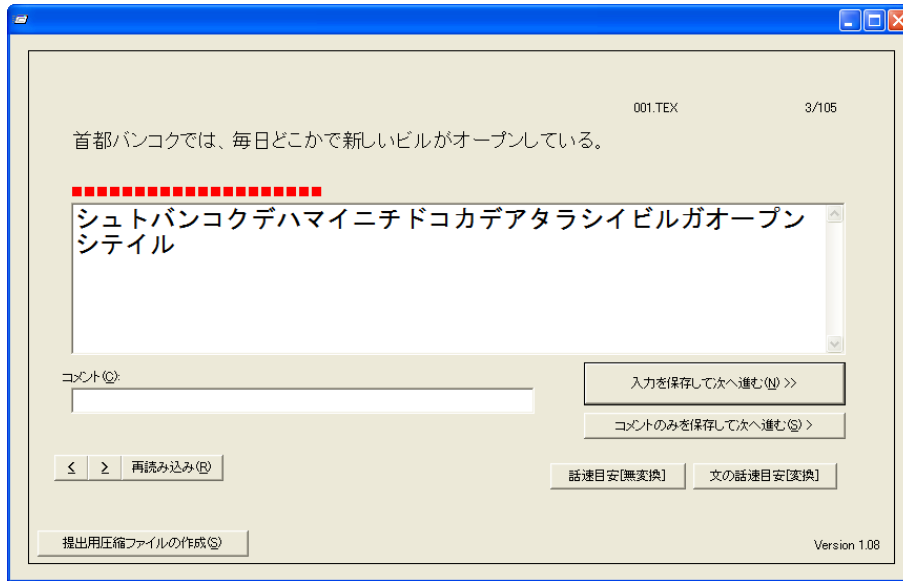


図 4.2: ラベリング作業用プログラムによる文発声ラベリング

作業用に、図 4.2 に示したプログラムを作成し、単純な操作でラベリング作業ができるようにした。ラベラは自宅のパーソナルコンピュータなどでこのプログラムを実行してラベリング作業をし、保存される結果を電子メールによって提出する。

上部にはラベリングの対象となる文の漢字仮名混じり表記が表示される。中央部の大きな入力欄（テキストボックス）には、読みが片仮名で表示され、この欄でラベリング作業を行なう。漢字仮名混じり表記の表示部と入力欄の間にある赤い色の四角形（ ）の表示は話速の目安を示すものであり、四角形は 1 モーラを表す。約 7 モーラ/秒の話速を提示するため、150 ミリ秒ごとに四角形が 1 つずつ増え、20 個に達したところで表示が消える。話速の目安は、新たな文を表示する度に必ず 1 回提示され、必要に応じてラベラ自身の操作によって表示することもできる。

入力欄はキーボードからの通常の入力を受け付けなくなっており、スペースキーを押すことで、カーソルの位置が [ ] (ラベル無し) [ ' ] (アクセント核を示すラベル) [ / ] (アクセント句境界を示すラベル) [ ] と変化する仕組みになっている。カーソルの位置は、矢印キーまたはマウス操作によって動かせるが、「キャ」「ミョ」などのように複数文字で 1 モーラを構成するものの内部（キとヤの間など）にはカーソルが行かないようになっている。また、促音「ッ」の付近にアクセント核を感じた場合にはその「ッ」の前をアクセント核とするのが一般的であるため、「ッ」の直後にはアクセント核ラベルを入力できない（アクセント句ラベルのみ入力できる）。いずれのラベルも文頭や文末には入力できず、複数のラベルを同じ位置に続けて配置することはいずれの位置でも不可能となっている。末尾にアクセント核がある形態素がアクセント句の末尾にある場合（単独でアクセント句を構成している場合を含む）にはアクセント句の末尾に潜在的なアクセント核が存在していることとなるが、ラベラにとって、そのようなアクセント核の存在を正確にラベリングすることは不可能であり、そもそも実際の現象としてアクセント核の役割が果たされていない以上、ラベリングの対象としないことが適切であろう。

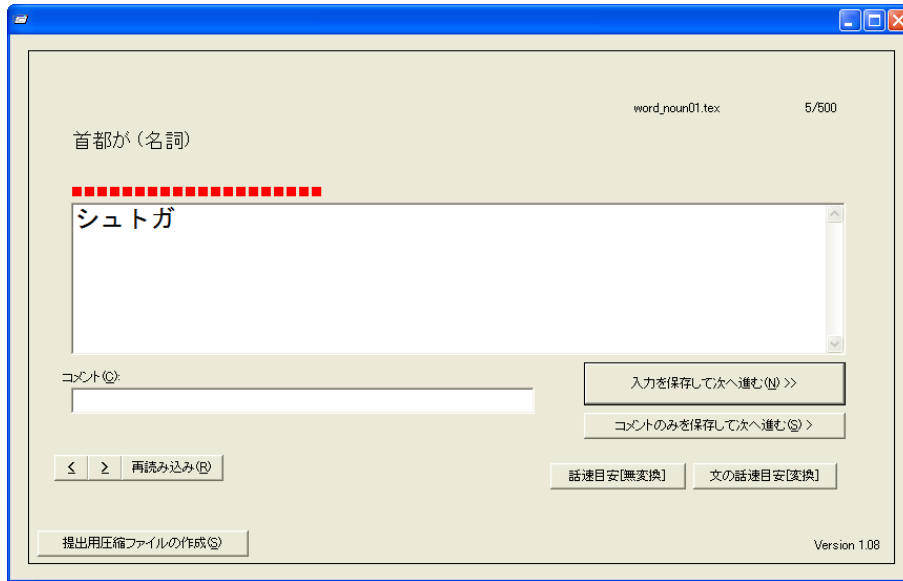


図 4.3: 形態素単独発声ラベリング

文に不明な点がある場合や表示された読み仮名が不自然であるなど、ラベラが何らかのコメントを記入する必要がある場合のための欄が、下部に設けられている。

自立語のラベリングでは、形態素を単独で発声する際のアクセントに関するのデータを採用することが目的であるが、完全に単独での発声では形態素が本来持っているアクセント情報を的確に保存されておらず曖昧性が残り、採取に適切でない。例えば、

- 名詞「妹」「大学」は、後ろに助詞「が」を付けて発声する場合には「イモ<sup>ー</sup>トガ」「ダイガクガ」のように異なるアクセントになるため、「妹」と「大学」は異なるアクセントの型を持っているといえる。しかし、単独で発声すると「イモ<sup>ー</sup>ト」「ダイガク」のように同じアクセントになってしまう。
- 形容詞「厳しい」「優しい」は後ろに名詞「もの」などが続く場合には「キビシ<sup>イ</sup>」「ヤサシ<sup>イ</sup>」のように明確に異なるアクセントで発声されるが、単独での発声では、「優しい」を「ヤサシ<sup>イ</sup>」することが許容される傾向にある。

があげられる。

以上に述べた曖昧性が残らないラベリングを行なうため、名詞には助詞「が」を後続させた状態でラベリングをさせ、形容詞には名詞「こと」が後続することを想定してラベリングをさせる（「こと」自体はラベリングの対象としない）などの表 4.1 に示す工夫を行なった。

この作業により、

- シュ 'ト / バ 'ンコクデ 'ハ / マ 'イニチノド 'コカデ / アタラシ 'イ / ビ 'ルガ / オ 'ーブ  
ン / シテイル

のような文発声ラベリングと、

- シュ 'トガ
- アタラシ 'イ

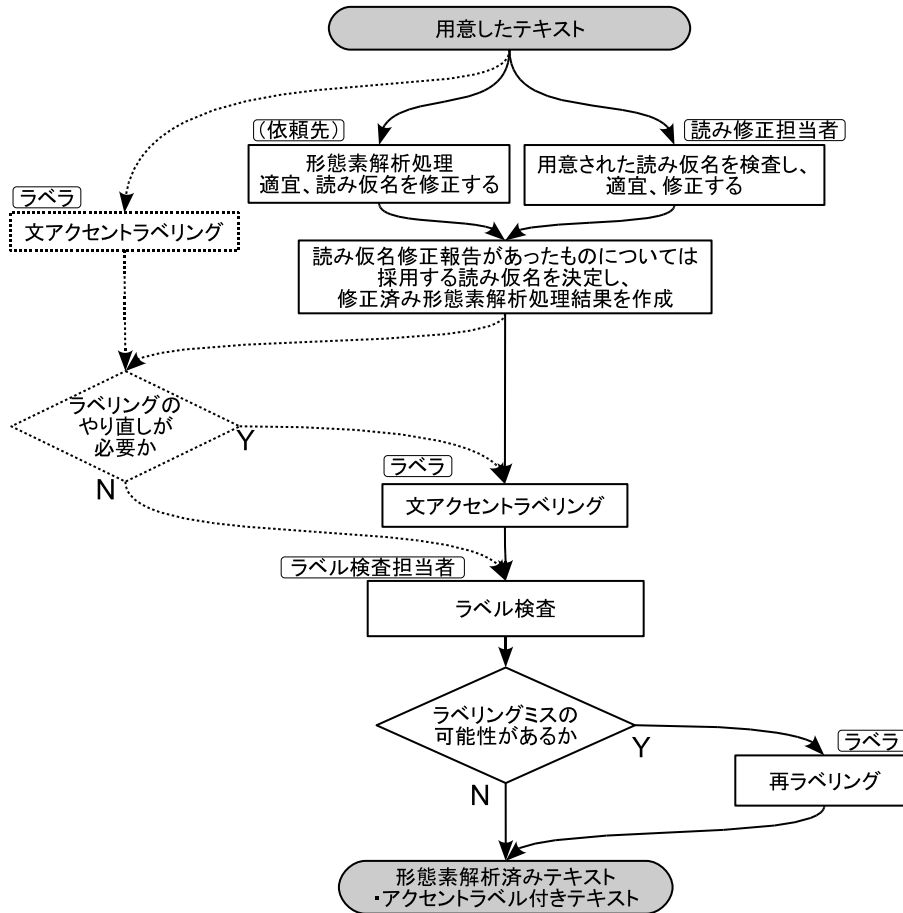


図 4.4: ラベリング作業のフローチャート（文データの流れ）

表 4.1: 形態素単独発声ラベリングでの形態素の提示方法

品詞	漢字仮名欄	品詞表示	入力欄
名詞-普通名詞-形状詞可能-*	～な	“名詞, 形容動詞”	～ナ
名詞-普通名詞-副詞可能	～が	“名詞, 副詞”	～ガ
名詞-固有名詞-*	～が	細分類まですべて	～ガ
名詞-数詞	～が	“数詞”	～ガ
名詞-*	～が	大分類のみ (“名詞”)	～ガ
代名詞	～が	大分類のみ (“代名詞”)	～ガ
形状詞-タリ	～と	“形容動詞-タリ”	～ト
形状詞-*	～な	“形容動詞”	～ナ
形容詞-*	～(こと)	大分類のみ (“形容詞”)	～
*	～	大分類のみ	～

- スル

のような形態素単独発声ラベリングが得られる。なお、[ / ] がアクセント句境界位置を、[ ' ] がアクセント核位置を表している。よって、「シュ 'ト」は「シト」を表している。

#### 4.2.9 ラベル検査作業・再ラベリング作業

作業効率を重視し、ラベラには素早く多くのラベリングをするよう求めた。従って、すべての文に対して誤り無くラベルを付与できたかどうかをラベラ自身が確認をしないまま、ラベリング結果が提出される可能性がある。そのため、ラベル検査担当者を用意し、ラベラが誤ったラベリングをした可能性がある文を指摘させ、指摘があった文のみ改めてラベラに丁寧に確認および訂正させることで、ラベリング結果の質を高めることとした。

ラベリング検査作業およびラベラによる確認訂正作業には、ラベリング作業とほぼ同様のプログラムを使用させた。ラベルが与えられていない読みが表示される代わりに、ラベラによるアクセントラベルの付いた読みが表示される点など何箇所か機能に差を付けており、これらの作業をしやすくように工夫してある。

### 4.3 完成状況と分析

#### 4.3.1 完成状況

本論文執筆の時点で、4,166 文 ( JNAS 新聞記事文の先頭 40 ファイル ) について、文発声・形態素単独発声の双方のラベリングが完了している。今後、用意したすべての文に対してのラベリングを目指して進めていく。

#### 4.3.2 分析

現時点で完成している 4,166 文での、アクセント句を構成する形態素の数を示したものが表 4.2 である。4 形態素以下のアクセント句が 90% 以上を占めていることが

アクセント句を構成する品詞列のうち出現率が 1% 以上のものを表にしたものが表 4.3 である。ここでの品詞は、形態素解析結果の品詞のうち、大分類のみを用いたものであり、品詞列は 8 位までのものの合計がおおよそ半数にあたることがわかり、また、出現率が 1% を下回る低い出現率のものの合計が 35% に達することなどが読み取れる。

表 4.2: アクセント句を構成する形態素の数

形態素数	出現数	出現率	累積
1	5079	17.4%	17.4%
2	9829	33.6%	51.0%
3	7902	27.0%	78.0%
4	3972	13.6%	91.6%
5	1586	5.4%	97.1%
6	554	1.9%	99.0%
7	200	0.7%	99.6%
8	69	0.2%	99.9%
9	24	0.1%	100.0%
10	6	0.0%	100.0%
11	3	0.0%	100.0%
12	1	0.0%	100.0%

表 4.3: アクセント句を構成する品詞列 (出現率 1%以上)

品詞列	出現数	出現率	累積
[名][助]	5273	18.0%	18.0%
[名]	2639	9.0%	27.1%
[名][名][助]	2180	7.5%	34.5%
[名][接尾][助]	1409	4.8%	39.4%
[動][助動]	792	2.7%	42.1%
[動]	788	2.7%	44.8%
[名][名]	758	2.6%	47.4%
[名][接尾]	739	2.5%	49.9%
[動][助]	571	2.0%	51.8%
[名][助][助]	541	1.9%	53.7%
[副]	483	1.7%	55.3%
[形容]	437	1.5%	56.8%
[名][名][接尾][助]	378	1.3%	58.1%
[名][助][名][助]	367	1.3%	59.4%
[代名][助]	334	1.1%	60.5%
[名][助][動]	316	1.1%	61.6%
[名][助動]	301	1.0%	62.6%
[名][動][助動]	297	1.0%	63.7%
[名][名][名][助]	296	1.0%	64.7%
上記以外	10326	35.3%	100.0%



# 第5章 統計的手法を用いたアクセント処理

## 5.1 統計的手法の導入

第3章で述べたように、規則に基づいたアクセント処理には限界があり、改善を必要とする。第2章で紹介した長野らの手法では  $n$ -gram を使用した統計的処理によって単語境界の推定と同時にアクセントを推定することに成功し、高い正答率を得ていたが、使用されたデータの詳細が不明であった。

本章では、第4章で作成したアクセントデータベースを使用しておこなった、条件付確率場 (CRF) に基づく統計的処理について述べる。

## 5.2 条件付確率場

### 5.2.1 条件付確率場

近年、注目されている統計的学習モデルに条件付確率場 (CRF; Conditional Random Fields) [21] がある。観測データ  $x$  を与えられた場合の出力ラベル  $y$  を学習する (図 5.1) にあたり、まず、CRF では、観測データと出力ラベルの対 ( $y_t$  と  $x_t$ ; 図 5.2) もしくは隣り合う出力ラベルの対 ( $y_{t-1}$  と  $y_t$ ; 図 5.3) についての特徴 (“素性” と呼ばれる)  $f$  として考えられるものを列挙する。例えば、「 $y_t$  が “noun (名詞) である” かつ  $x_t$  が “大文字で始まる”」「 $y_{t-1}$  が “noun (名詞) である” かつ  $y_t$  が “verb (動詞) である”」などがこれに該当する。なお、観測データと出力ラベルの対に関する素性を観測素性、隣り合う出力ラベルの対に関する素性を遷移素性と呼ぶ。その上で、各素性  $f$  について重要度  $\theta_f$  (負の値にもなりうる) を考え、 $(x, y)$  内で素性  $f$  が満たされている箇所の数を  $\phi_f(x, y)$  とおく。入力  $x$  に出力  $y$  を割り当てること確信度合いは、素性が満たされる度にその素性の重要度を足し合

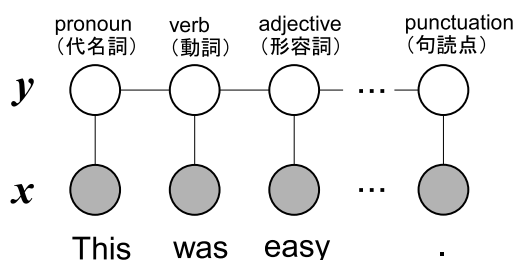


図 5.1: 観測データ ( $x$ ) と出力ラベル ( $y$ )

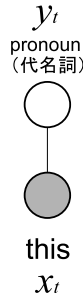


図 5.2: 観測素性

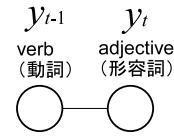


図 5.3: 遷移素性

わせたものとして考えることができ、 $\sum_f \theta_f \phi_f(\mathbf{x}, \mathbf{y})$  と表すことができる。これを、

$$P(\mathbf{y}|\mathbf{x}) = \frac{\exp \sum_f \theta_f \phi_f(\mathbf{x}, \mathbf{y})}{\sum_{\mathbf{y}} (\exp \sum_f \theta_f \phi_f(\mathbf{x}, \mathbf{y}))}$$

とすることで0以上1以下の値をとる確率分布とすることができ、 $x$  に対して  $y$  のラベルが付与される確率を表す値として用いることができる。正解ラベルデータを与えることによる学習は、与えられた多くの  $x$  と  $y$  の組を使用した上で、この確率値  $P(\mathbf{y}|\mathbf{x})$  をできるだけ大きくするような重要度  $\theta_f$  を探る最尤推定処理、すなわち、

$$\mathcal{L}_{\theta} = \sum_i \log(P(\mathbf{y}_i|\mathbf{x}_i))$$

をできるだけ大きくする処理となる。そして、与えられた  $x$  から  $y$  を求めるラベル推定は、 $x$  と学習の段階で決定された重要度  $\theta_f$  と、最も  $P(\mathbf{y}|\mathbf{x})$  が高くなる  $y$  を見つける処理となる。

学習での、適切な重要度  $\theta_f$  を求める最尤推定では、動的計画法を用いて工夫がされ、効率的に解を求められるように処理がなされる。しかし、単純な最尤推定によって辿り着く解は、学習データに適応しすぎた過学習の状態に至ることが多いため、

$$\mathcal{L}_{\theta} = C \sum_i \log(P(\mathbf{y}_i|\mathbf{x}_i)) - \frac{1}{2} \sum_k |\theta_k|$$

などとした上で、 $C$  (ハイパーパラメータ) の大きさを調整することにより学習データへの適応度を变化させる。

## 5.2.2 CRF++

本研究では、CRF による処理を行なうツールキットとして CRF++[23] を使用した。CRF++は、スペースまたはタブで区切られたテキストファイルを学習・推定用のデータとして用いる。

学習に使用する素性については、テンプレートと呼ばれるファイルを用意し、その中に記述する。入力する学習・推定データは、それぞれの観測データ  $x_t$  に関しての使用しうる特徴

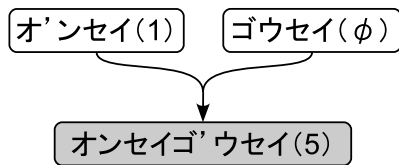


図 5.4: アクセント結合

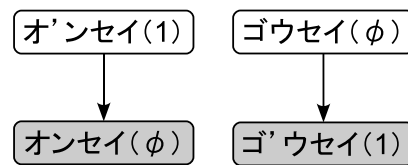


図 5.5: アクセント結合を個々の形態素のアクセント変化と捉える場合

をタブ区切り（または空白区切り）で列挙されたものとして用意する。テンプレートファイルには何番目の列の特徴を観測素性で用いるかについて記入しておけば、プログラムの実行時に特徴がすべて展開される。例えば、学習データの各行 2 列目に「その語が大文字で始まるか」という特徴（値は“Y”か“N”であるとする）が記入されている場合、テンプレートに 2 列目を学習に使用する旨が記されていれば、「その語が大文字で始まるかが“Y”であり出力ラベルが“noun”である」「その語が大文字で始まるかが“N”であり出力ラベルが“verb”である」など、2 列目に出現しうる特徴と出力ラベルのすべての組が観測素性として用意される。特徴としては、着目している行以外に記されているものを使用することもでき、例えば、テンプレートに着目している語の前の語の 2 列目を学習に使用する旨を記すと、「その語の前の語が大文字で始まるかが“Y”であり...」などの観測素性が用意される。先頭の語の前にある語を参照しようとした場合などのように、参照先が存在しない場合には、“B-1”（先頭の 1 行前），“B+2”（末尾の 2 行後）などの特殊な値が参照先の値として返される。また、遷移素性については、遷移素性を使用する旨の記述をしておけば、出現するすべての出力ラベル 2 つについての遷移素性が自動的に用意される。

なお、学習の際には前述のハイパーパラメータ  $C$  を指定する必要があり、以降で示す実験では、すべて  $C = 10$  とした。類似したデータで事前に  $C$  を数種類試したところ、おおむね適切な結果が得られた値として採用したものである。

## 5.3 条件付確率場を用いたアクセント処理の学習・推定

### 5.3.1 学習・推定の内容

この章では、テキストコーパスとして、第 4 章で作成したアクセントデータベースの完成分 4,166 文を使用した、アクセントの統計的学習・推定について述べる。形態素解析結果や自立語単独発声時のアクセントについても、同データベースのものを使用した。

ここでは、形態素解析とアクセント句境界の配置が既に正しく行なわれたデータがある状況を想定し、各アクセント句内のアクセントを正確に処理することを目的とする。すなわち、今回の実験では、アクセントデータベースに含まれる、形態素解析データとラベラによるアクセント句境界位置のデータは常に既知のものとして扱い、ラベラによるアクセント核位置のデータを学習・推定の対象とする。

実際に学習と推定を行なうにあたり、データベースに含まれるアクセント句境界位置情報を使用して文をアクセント句に区切り、句に含まれる各形態素の文中アクセント型を学習と推定の対象とした。ここでいう文中アクセント型とは、「音声合成」であれば「オンセー」

	A	B	C	D	E	F	G	H	I	J	K
1	まだ	マダ	未だ	マダ	副詞	*	*	マ'ダ/	2	1	1
2											
3	正式	セイシキ	正式	セイシキ	名詞-普通名詞-形状詞可能	*	*	セイシキ	4	-1	-1
4	に	ニ	だ	ダ	助動詞	助動詞-ダ	連用形-二	ニ/	1*	-1	
5											
6	決まっ	キマッ	決まる	キマル	動詞-一般	五段-ラ行-一般	連用形-促音便	キマッ	3	-1	-1
7	た	タ	た	タ	助動詞	助動詞-タ	基本形-一般	タ/	1*	-1	
8											
9	わけ	ワケ	訳	ワケ	名詞-普通名詞-一般	*	*	ワ'ケ	2	1	1
10	で	デ	で	デ	助詞-格助詞	*	*	デ'	1*	1	
11	は	ハ	は	ハ	助詞-係助詞	*	*	ハ/	1*	-1	
12											
13	ない	ナイ	無い	ナイ	形容詞-非自立可能	形容詞-ア段-無イ+ない	基本形-一般	ナ'イ	2	1	1
14	の	ノ	の	ノ	助詞-準体助詞	*	*	ノ'	1*	1	
15	で	デ	だ	ダ	助動詞	助動詞-ダ	連用形-一般	デ	1*	-1	
16											
17	カネ	カネ	金	カネ	名詞-普通名詞-一般	*	*	カネ	2	-1	-1
18	の	ノ	の	ノ	助詞-格助詞	*	*	ノ'	1*	-1	
19	力	チカラ	力	チカラ	名詞-普通名詞-一般	*	*	チカラ'	3	3	3
20	も	モ	も	モ	助詞-係助詞	*	*	モ/	1*	-1	
21											
22	十分	ジュウブン	十分	ジュウブン	副詞	*	*	ジュウブ'ン/	4	3	3
23											
24	知っ	シッ	知る	シル	動詞-一般	五段-ラ行-一般	連用形-促音便	シッ	2	-1	-1
25	て	テ	て	テ	助詞-接続助詞	*	*	テ'	1*	-1	
26	い	イ	居る	イル	動詞-非自立可能	上一段-ア行	連用形-一般	イ	1	-1	-1
27	た	タ	た	タ	助動詞	助動詞-タ	基本形-一般	タ	1*	-1	
28											
29	首都	シュト	首都	シュト	名詞-普通名詞-一般	*	*	シュ'ト/	2	1	1
30											
31	バンコク	バンコク	バンコク	バンコク	名詞-固有名詞-地名-一般	*	*	バン'コク	4	1	1
32	で	デ	で	デ	助詞-格助詞	*	*	デ'	1*	1	
33	は	ハ	は	ハ	助詞-係助詞	*	*	ハ/	1*	-1	

図 5.6: 形態素解析結果と対応付けたアクセントラベル

(無核)と「ゴーセー」(1モーラ目に核)を指す。すなわち、「音声(オンセー)」と「合成(ゴーセー)」から「音声合成(オンセーゴーセー)」が作られるアクセント結合現象(図 5.4)を、形態素「音声」「合成」のそれぞれのアクセント型が変化した(単独発声アクセント型とは異なる文中アクセント型になった; 図 5.5)ものと見做して扱う。文中で形態素の境界にアクセント核が生じているもの(「流れ・を(ナガレ・オ)」「出来・ない(デキ・ナイ)」など)では、当該アクセント核は前側の形態素に属するものとした。

このようにして形態素解析結果およびラベラによるアクセント句区切り・文中アクセント型・単独発声アクセント型を組み合わせるとまとめた結果が図 5.6 である。列 H が文中アクセントラベル済みテキストを形態素に対応付けたもの、列 K が形態素の文中アクセント型、列 J が形態素の単独発声アクセント型である。なお、表中の“-1”はアクセント核が存在しないことを表し、“\*”は単独アクセント型ラベリングの対象となっていない形態素であることを表す。その他の数値は核のあるモーラの位置を表している。なお、図 5.6 に掲げられている各形態素の属性は、主要なものを抜き出したものであり、実際にはより多くを持つ。

なお、この過程で、形態素内にアクセント句境界が存在している文、複数のアクセント核を持つ形態素を含む文は文ごと除去した結果、使用できる文は全部で 4,108 文となった。学習・推定には、これらの文を学習用 3,581 文(25,692 アクセント句)、推定用 527 文(3,533 アクセント句)に分けて使用した。この区分けは、データベースの基となる JNAS 新聞記事文の 35 ファイルと 5 ファイルとして分割したものである。

これらのデータを用いて、

- 統計的手法による文中アクセント型の学習・推定

表 5.1: 統計処理による正答数（単独型なし、当該形態素の特徴のみ使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	8498	85.8%	—	—%
アクセント句単位	すべての句	3533	2403	68.0%	2476	70.1%
	単純な句	822	674	82.0%	676	82.2%
	名詞連続を含む句	688	382	55.5%	390	56.7%

表 5.2: 統計処理による正答数（単独型なし、前後 1 形態素までの特徴を使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9054	91.4%	—	—%
アクセント句単位	すべての句	3533	2813	79.6%	2884	81.6%
	単純な句	822	714	86.9%	718	87.3%
	名詞連続を含む句	688	497	72.2%	507	73.7%

表 5.3: 統計処理による正答数（単独型なし、前後 2 形態素までの特徴を使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9080	91.6%	—	—%
アクセント句単位	すべての句	3533	2833	80.2%	2899	82.1%
	単純な句	822	703	85.5%	706	85.9%
	名詞連続を含む句	688	521	75.7%	530	77.0%

- 単独発声アクセント型を与えない場合（5.3.2）
- 単独発声アクセント型を与える場合（5.3.3）
- 規則に基づく文中アクセント型の推定（5.3.4）

の 3 種類の文中アクセント推定を行ない、比較をする。

### 5.3.2 単独発声アクセント型を与えない場合における文中アクセント型の学習・推定

前節のデータを使用して、文中アクセント型の学習と推定を行なった。まず、形態素単独発声アクセント型が与えられていない状況を想定するものとし、

- 観測素性
  - 当該形態素および前後  $n$  形態素のそれぞれ ( $n \leq 2$ ) について、
    - [基本形, 基本形読み, 書字形, 品詞, 活用法] (組み合わせ; 形態素の同定に用いる)

- [品詞]
- [品詞 (大分類のみ)]
- [活用型 (大分類のみ)]
- [活用形 (大分類のみ)]
- [モーラ数]

と当該形態素の文中アクセント型との関係

- 遷移素性 (隣接する形態素の文中アクセント型同士の関係)

を素性として指定した。 $n$  の値として 0,1,2 の 3 種類、つまり、当該形態素の情報のみを用いるもの、前後 1 形態素を含む 3 形態素の情報を用いるもの、前後 2 形態素を含む 5 形態素の情報を用いるものの 3 種類である。 $n = 0$  とした時の結果が表 5.1、 $n = 1$  では表 5.2、 $n = 2$  では表 5.3 である。

各表は、「形態素単位」とした行は、形態素ごとに推定結果とデータベース中の正解データが一致したものの数および割合を示したものである。「アクセント句単位」はアクセント句としての一致の数と割合を示したものであるが、このうち、「すべての核」とした列はアクセント句内のすべての形態素のアクセントが一致したもの、すなわち、アクセント句内に複数のアクセント核が存在する場合にはそれらがすべて一致した場合にのみ正答として集計したものである。それに対し、「主核のみ」としたものは、アクセント句内の 2 つめ以降のアクセント核は無視することとし、推定結果と正解において 1 つめのアクセント核が一致した場合または双方がアクセント核を持たない場合に正答と認めて集計したものである。

また、アクセント句の構成によって正答率にばらつきがあるため、特に高い正答率となる「単純なアクセント句 ( { 名詞, 動詞, 形容詞, 形状詞 } + { 助詞, 助動詞 } の 2 語で構成されたもの)」および特に低い正答率となる「名詞の連続を含むアクセント句」は該当するアクセント句のみについても集計した。

### 5.3.3 単独発声アクセント型を与える場合における文中アクセント型の直接学習・推定

次に、形態素単独発声アクセント型が与えられている状況を想定し、前節での素性に加え、観測素性に

- - [単独発声アクセント型]

を加えて同様の学習を行なった。当該形態素の情報のみを用いるもの ( $n = 0$ ) の結果が表 5.4、前後 1 形態素を含む 3 形態素の情報を用いるものでは表 5.5、前後 2 形態素を含む 5 形態素の情報を用いるものでは表 5.6 である。

### 5.3.4 規則に基づいた文中アクセント型の推定

統計的処理との比較のために、アクセント結合規則によるアクセント処理を行なった。

規則に基づく処理には、GalateaTalk に含まれるアクセント処理モジュールを単体で使用した。ただし、使用の前に以下の点についての処理を書き加えた。

表 5.4: 統計処理による正答数（単独型あり，当該形態素の特徴のみ使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	8638	87.2%	—	—%
アクセント句単位	すべての句	3533	2535	71.8%	2605	73.7%
	単純な句	822	736	89.5%	738	89.8%
	名詞連続を含む句	688	376	54.7%	382	55.5%

表 5.5: 統計処理による正答数（単独型あり，前後 1 形態素までの特徴を使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9229	93.1%	—	—%
アクセント句単位	すべての句	3533	2975	84.2%	3054	86.4%
	単純な句	822	776	94.4%	782	95.1%
	名詞連続を含む句	688	501	72.8%	510	74.1%

表 5.6: 統計処理による正答数（単独型あり，前後 2 形態素までの特徴を使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9272	93.6%	—	—%
アクセント句単位	すべての句	3533	3004	85.0%	3081	87.2%
	単純な句	822	769	93.6%	775	94.3%
	名詞連続を含む句	688	511	74.3%	523	76.0%

表 5.7: 規則に基づく手法の正答数

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		—	—	—%	—	—%
アクセント句単位	すべての句	3430	2619	76.4%	2633	76.8%
	単純な句	822	776	94.4%	777	94.5%
	複合名詞を含む句	675	496	73.5%	501	74.2%

- 第3章で述べた、活用によるアクセント変化。
- 田中の「式」(2.4節)の「従接式」に相当する処理。表3.1に示したように、アクセント結合様式 F1,F2 に属する付属語は「従接式」「独立式」のどちらに属するか判断できないが、すべてを出現事例の多い「従接式」と見做した。
- 文節間結合規則(表2.4)の実装。GalateaTalk にはおよそ文節に相当するものをアクセント句として扱うため、文節間結合規則が実装されていない。本実験ではアクセント句を外部から与えるため、通常の GalateaTalk では複数のアクセント句として出力される入力、文節間結合規則を使用して必ず1つのアクセント句にする。

以上のようなモジュールに、推定用527文(3,533アクセント句)を入力して、各アクセント句中の文中アクセント型を出力させた。なお、各形態素の単独発声アクセント型やアクセント結合型(アクセント様式・結合アクセント価)は入力するデータとして与えた。単独発声アクセント型は、ラベラによる単独発声ラベルを使用し、アクセント結合型については形態素解析結果のデータに含まれるもの(UniDicによるもの)を使用した。この結果が、表5.7である。処理結果はアクセント句単位で出力されるため、形態素単位での正答数・正答率については不明である。また、アクセント句総数が表5.1などのものより少ない値となっているが、これは、アクセント処理モジュールが対応しておらず結果を出力できないアクセント句があったため、それらを母集団から外したことによるものである。これらのアクセント句を、誤答を出力したものとみなすこととすると、表5.7よりも若干正答率が低いものとなる。

### 5.3.5 考察

条件付確率場を使用した学習・推定で、観測素性に当該形態素の特徴のみを使用したもの(表5.1,表5.4)は、どのような場合について見ても、規則に基づく処理(表5.7)に比べて正答率が低い。形態素が文中でいかなるアクセントは、当該形態素のみの特徴だけで決まるものではなく、周囲の形態素などとの関係において決定されるものであることが現れていると考えられる。

それに対し、前後1形態素または2形態素についての特徴を用いた場合には、概ね、規則に基づく処理より優れた結果が得られている。形態素の単独発声アクセントについての情報が与えられない環境であっても、前後1形態素までの特徴を用いて学習・推定をすれば(表5.2)、主アクセント核のみについて考える場合で81.6%のアクセント句について正しい文中アクセントが得られ、規則に基づく処理での76.8%を大きく上回っている。また、2.5節で紹介した長野らの手法と比較し、学習に使用したデータの量が半分に満たないにもかかわらず、形態素単位でのアクセント推定率がわずかに劣るにとどまっておらず、高い正答率が得られていると言える。

しかし、単純なアクセント句のみに限って考えると、規則での94.5%にまったく及ばない87.3%にとどまっている。従って、単純なもの以外のアクセント句の処理について規則が不十分であることが示されたと考えられるべきであり、単独発声アクセント型が与えられない状況での統計的処理が優れた結果を出せるわけではないと言える。

単独発声のアクセントを与えた場合の結果について見ると、前後1形態素または2形態素



についての特徴を用いた場合（表 5.5, 表 5.6）には、ほぼいかなる場合においても規則に基づく処理（表 5.7）を上回る正解率を得ている。

用いる特徴を前後 1 形態素までとする場合（表 5.2, 表 5.5）と 2 形態素までとする場合（表 5.3, 表 5.6）の比較では、すべての句についての正答率や名詞連続を含む句についての正答率では前後 2 形態素までの特徴を用いた方が高い率を得ているが、単純な結合のみに限ってみれば、前後 1 形態素までの特徴を用いたものの方が正答率が高いが、単純な結合として扱っているアクセント句はすべて 2 形態素のみからなるものであるため、前後 2 形態素までを見ることによる利点はないのが当然であり、学習データに対する過学習としての欠点のみが現れたものであると言える。

## 5.4 アクセント結合規則を参考とした学習・推定の工夫

ここまでの実験では、単独発声アクセント型を与えた上で前後 2 形態素までの特徴を用いて学習・推定をすると、最も良い正答率を得た。本節では、規則で行なわれる処理を参考に、学習の方法を改善し、正答率の向上を目指した。

### 5.4.1 単独発声アクセント型から文中アクセント型への変化についての学習・推定

句坂らのアクセント結合規則を見て分かるように、アクセント結合で結合前のアクセント句の位置が保存されることは多い。前節の手法では、文中アクセント型を直接学習や推定の対象としていたため、単独発声型と文中アクセント型がいずれも“1”である場合といずれも“2”である場合は別々の事象として扱われることとなる。そこで、単独発声アクセント型から文中アクセント型への変化を学習・推定の対象とすることで、類似する現象を共通のものと捉えられるようにし、学習・推定をした。具体的には、

- 単独発声アクセント型が有核のもの（“-1”以外の値を持つもの）に対しては、
  - 文中アクセント型が有核の場合、単独発声型からの変化量を表すラベル（“[0]”; “[+1]”, “[+2]”, ...; “[-1]”, “[-2]”, ...）
  - 文中アクセント型が無核（“-1”）の場合、無核を示すラベル（“non”）
- 単独発声アクセント型が無核のもの（“-1”）または値を持たないものに対しては、
  - 文中アクセント型（“1”, “2”, ...; “-1”）

を機械的な処理によって与え、そのラベル（相対変化ラベルとする）を学習・推定の対象とした。素性には、前後 2 形態素についての 5.3.3 節における実験と同様の特徴を用いて行なった。推定結果としては相対変化ラベルが得られることとなるため、機械的に文中アクセント型を表す数値に復元し、正答数の算出した。その結果が、表 5.8 である。文中アクセント型を直接学習するもの（表 5.6）と比較し、すべての場合において正答率が向上している。相対的な変化を学習の対象とすることで、効率の良い学習ができたものと言える。

表 5.8: 相対変化ラベルの学習による正答数 (前後 2 形態素までの特徴を使用)

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9319	94.1%	—	—%
アクセント句単位	すべての句	3533	3050	86.3%	3137	88.8%
	単純な句	822	788	95.9%	791	96.2%
	名詞連続を含む句	688	540	78.5%	553	80.4%

#### 5.4.2 隣接形態素組み合わせ素性を用いた学習・推定

ここまでを用いた観測素性は、当該形態素または周辺形態素の品詞等の情報 1 つと出力ラベルの関係のみであった。結果として、句坂らによる複合単語アクセント結合規則のように名詞が連続した場合にのみ起こる現象を学習する場合、例えば、表 2.2 の様式 C1 のように当該形態素が名詞で 1 つ後の形態素も名詞である場合に文中アクセント型が無核になることを学習する場合、2 つの観測素性

- 当該形態素の品詞が名詞である場合には無核になる
- 1 つ後の形態素の品詞が名詞である場合には無核になる

についての重要度  $\theta_f$  をあげる処理がされる。しかし、これは名詞の連続以外に対しても影響を与えるため、特定の組み合わせに対する学習には不向きである。そこで、これまでの、

- 当該形態素および前後 2 形態素のそれぞれについての、
  - [基本形, 基本形読み, 書字形, 品詞, 活用型] (組み合わせ; 形態素の同定に用いる)
  - [品詞]
  - [品詞 (大分類のみ)]
  - [活用型 (大分類のみ)]
  - [活用形 (大分類のみ)]
  - [モーラ数]
  - [単独発声アクセント型]

と当該形態素の文中アクセント型との関係

に加え、

- 隣接形態素の組み合わせとしての
  - [当該形態素の品詞, 前の形態素の品詞]
  - [当該形態素の品詞, 後の形態素の品詞]
  - [当該形態素の品詞, 前の形態素の { 基本形, 基本形読み, 書字形, 品詞, 活用型 }]
  - [当該形態素の品詞, 後の形態素の { 基本形, 基本形読み, 書字形, 品詞, 活用型 }]
  - [当該形態素の { 基本形, 基本形読み, 書字形, 品詞, 活用型 }, 前の形態素の品詞]

表 5.9: 相対変化ラベルの学習による正答数（前後と組み合わせた特徴も加えて使用）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9424	95.1%	—	—%
アクセント句単位	すべての句	3533	3129	88.6%	3214	91.0%
	単純な句	822	785	95.5%	790	96.1%
	名詞連続を含む句	688	566	82.3%	578	84.0%

－ [当該形態素の { 基本形, 基本形読み, 書字形, 品詞, 活字型 }, 後の形態素の品詞]

と当該形態素の文中アクセント型との関係

を観測素性として使用し、相対変化ラベルの学習・推定をした。品詞と品詞の組み合わせ以外にも、当該形態素の品詞と前後形態素の基本形等、当該形態素の基本形等と前後形態素の品詞について組み合わせることで、例外的なアクセントについても学習がされやすいように配慮した。その結果が表 5.9 である。名詞連続を含むアクセント句について特に顕著な改善が見られ、組み合わせの学習が効果を発揮していると言える。

#### 5.4.3 アクセント結合規則に適合させた学習・推定

以上のように、アクセントの変化を学習の対象とするなど、句坂らの規則で可能な処理と類似した結果が得られる学習を行なうことで、高い正答率が得られた。よって、規則で表現されている処理とさらに取り入れた形で学習をするにより、規則的現象をさらに的確に捉えることができるものと考えられる。

句坂らの規則によって得られる結果をまとめると、次のように分類できる。

- 単独発声時のアクセント核がそのまま残る
- 単独発声時のアクセント核を失う
- 単独発声時に核が無かったものがそのまま無核となる
- 末尾に核が現れる
- 1 モーラ目が核になる

また、結合前のアクセント型（単独発声アクセント型）によって、上記分類における異なる結果が選ばれるのは、以下のような場合である。

- 付属語規則 ..... 1 語目が無核のとき ( $M_1 = \phi$  のとき) とそうでないとき
- 複合名詞規則 ..... 2 語目が無核であるか最後の音節に核があるとき ( $M_2 = N_2$  (or  $N_2 - 1$ ),  $\phi$  のとき) とそうでないとき
- 接頭辞規則 ..... 2 語目が無核であるか最後の音節に核があるとき ( $M_2 = N_2$  (or  $N_2 - 1$ ),  $\phi$  のとき) とそうでないとき
- 文節間規則 ..... 1 語目が無核のとき ( $M_1 = \phi$  のとき) とそうでないとき

また、アクセント結合規則適用則は、以下のようなアクセント変化を生じさせるものと捉えられる。

- 巡回的適用則 …… (アクセント型の相対変化に関する規則ではない)
- 一段活用動詞処理規則 …… 未然形・連用形ではアクセント型の値が1減る
- 無声化に伴う移動規則 …… 無声化した母音が核を持つことになる場合はアクセント型の値が1減る
- 音節内移動則 …… 規則上アクセント核が置かれることとなるモーラが特定のもの(「ー」「ン」「ッ」などの特殊拍)の場合は1減る
- 先行品詞による結合アクセント価選択則 …… (アクセント型の相対変化に関する規則ではない)

よって、アクセント結合規則適用則によって、「単独発声時のアクセント核の1つ前が核になる」「末尾の1つ前が核になる」といった結果が導かれる場合があることがわかる。これらを総合的に考え、以下のような相対変化ラベルを導入し、これを学習・推定の対象とすることにした。

- 単独発声アクセント型が有核のもの(“-1”以外の値を持つもの)に対しては、
  - 文中アクセント型が無核(“-1”)の場合、その旨を示すラベル(“non”)
  - 文中アクセント型が単独発声型と同じである場合、その旨を示すラベル(“same”)
  - 文中アクセント型がモーラ数と同じである(末尾に核がある)場合、その旨を示すラベル(“morae”)
  - 文中アクセント型が単独発声型より1小さい場合、その旨を示すラベル(“same-1”)
  - 文中アクセント型が1型の場合、その旨を示すラベル(“one”)
  - 文中アクセント型がモーラ数より1小さい(末尾の1つ前に核がある)場合、その旨を示すラベル(“morae-1”)
  - その他の場合は、単独発声型からの変化量を表すラベル(“[0]”; “[+1]”, “[+2]”, …; “[-1]”, “[-2]”, …)
- 単独発声アクセント型が無核のもの(“-1”)に対しては、
  - 文中アクセント型が単独型と同様に無核(“-1”)の場合、その旨を示すラベル(“samennon”)
  - 文中アクセント型がモーラ数と同じである(末尾に核がある)場合、その旨を示すラベル(“morae”)
  - 文中アクセント型が1型の場合、その旨を示すラベル(“one”)
  - 文中アクセント型がモーラ数より1小さい(末尾の1つ前に核がある)場合、その旨を示すラベル(“morae-1”)
  - その他の場合は、文中アクセント型(“1”, “2”, …; “-1”)
- 単独発声アクセント型を持たないものに対しては、
  - 文中アクセント型(“1”, “2”, …; “-1”)

表 5.10: 規則を参考とした変化の学習による正答数

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	9458	95.5%	—	—%
アクセント句単位	すべての句	3533	3159	89.4%	3238	91.7%
	単純な句	822	785	95.5%	792	96.4%
	名詞連続を含む句	688	574	83.4%	589	85.6%

これらのうち、複数に該当するアクセント句については、最も上にあるものを相対変化ラベルとして使用する。

また、学習・推定のデータに、

- 単独発声アクセント型が無核である場合、その旨を示すラベル (“non”)
- 単独発声アクセント型がモーラ数と同じである (末尾に核がある) 場合、その旨を示すラベル (“morae”)
- 単独発声モーラ数より 1 小さい (末尾の 1 つ前に核がある) 場合、基本形読みの末尾 2 モーラ (“オイ” など)
- 単独発声型が上記に当てはまらない場合、その旨を示すラベル (“else”)

とするラベル (“単独型種類ラベル” とする) を用意し、単独発声アクセント型の種類による結果の分岐を行ないやすくすることを目指した。

また、学習・推定の際に使用する観測素性に以下のものを加えた。

- 当該形態素についての、
  - [単独型種類ラベル]
  - [出現形読みの先頭のモーラ]
  - [出現形読みの第 2 モーラ]
  - [出現形読みの、単独発声時のアクセント核位置の 1 つ前に位置するモーラ]
  - [出現形読みの、単独発声時のアクセント核位置のモーラ]
  - [出現形読みの、単独発声時のアクセント核位置の 1 つ後に位置するモーラ]
  - [出現形読みの末尾の 1 つ前に位置するモーラ]
  - [出現形読みの末尾のモーラ]

と当該形態素の文中アクセント型との関係

出現形の読みの各モーラを用いたのは、音節内移動則などの影響を学習するためである。

上記のように、出力ラベルや観測素性に工夫をして学習・推定をした結果が、表 5.10 である。すべての場合に渡って、表 5.9 からの正答率の向上が見られる。

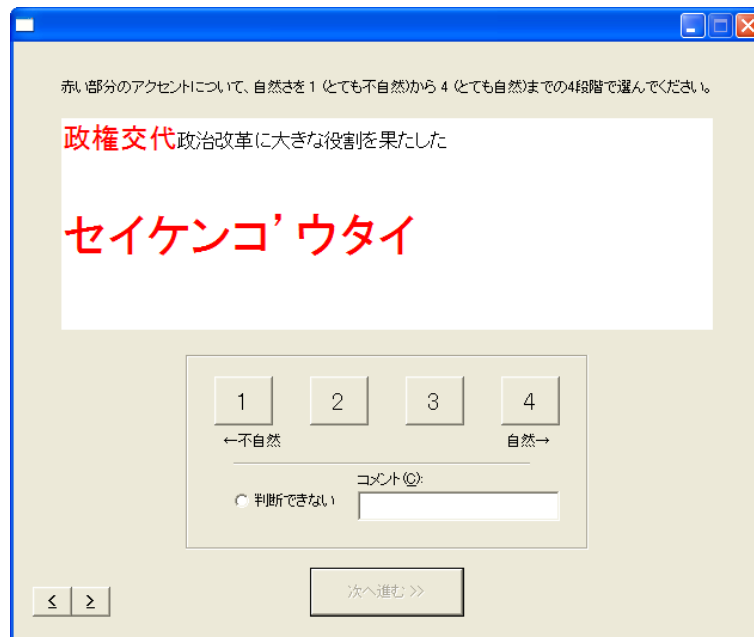


図 5.7: 許容度判定プログラム

#### 5.4.4 不正解に分類された形態素の許容度に関する調査

統計的なアクセント学習に各種の工夫を重ねることにより、アクセント推定は、91.7%のアクセント句に対して正しい結果が得られており（主アクセント核のみについて考える場合）単純なアクセント句に限れば正答率は96.4%に及ぶ。しかし、このことは、単純なアクセント句であっても3.6%の句ではアクセント推定結果と正解ラベルが一致しないことを意味し、それらの不一致が本質的な誤りであるか、アクセントの揺れの範囲内として許容できるものなのかによって、結果の意味するところは異なる。そこで、アクセント推定結果と正解ラベルが一致しなかったアクセント句について、第4章でアクセントラベリングを行なったラベラによる許容度の判定を行なった。

許容度の判定には図5.7のプログラムを使用し、5.4.3節の実験において正答とならなかった（主アクセント核のみについて考える場合にだけ正答となったものを含む）アクセント句374句（= 3533 - 3159）に対して、1（とても不自然）から4（とても自然）の4段階で許容度を回答させ、3以上の値を許容と解釈することとした。結果が、表5.9である。なお、「回答不可」欄は、アクセント句が不自然であるとして回答がなされなかったものの数である。

また、許容度判定で許容されたアクセント句を正答と見做して表5.10を集計しなおしたものが、表5.12である。単純な結合では、98.3%ものアクセント句について正しい結果または許容できる結果が得られていることとなった。音声として出力する場合にはさらに許容できる幅が広がると考えられるため、完全に近い精度であると言えるだろう。ただし、許容できないと判定されたアクセント句がわずかながら存在したのは事実であり、何らかの方法で改善することが望まれる。

表 5.11: 許容度調査の結果

	対象数	回答された許容度					許容数
		1	2	3	4	回答不可	
すべての句	374	194	30	74	74	2	148
単純な句	37	13	1	7	16	0	23
名詞連続を含む句	114	74	8	19	12	1	31

表 5.12: 規則を参考とした変化の学習による正答数（許容される句も正答とする場合）

		総数	すべての核		主核のみ	
			正答数	正答率	正答数	正答率
形態素単位		9908	—	—%	—	—%
アクセント句単位	すべての句	3533	3307	93.6%	—	—%
	単純な句	822	808	98.3%	—	—%
	名詞連続を含む句	688	605	87.9%	—	—%

#### 5.4.5 考察

統計的な学習にアクセント結合規則から得られる知識を利用し、正答率を高めることに成功した。

隣接する形態素の品詞を組にしたものを観測素性に加えたことによる改善からは、句坂らのアクセント結合規則で結合する品詞ごとに規則を使い分けられているように、品詞がどのように連続するかが文中でのアクセントを考える上で重要であることを表しており、当該形態素や前後の形態素の品詞を個別に考えるのでは不十分となるほどに大きな影響を与えているということが分かる。

アクセント結合規則に細かく適合させた学習において正答率の向上が見られたことから、規則が完全であるとは言えないものの、アクセント結合処理の重要な部分を押さえていることが分かる。今回の実験では、アクセント結合規則への細かな適合をする前とした後を比較したため、どの部分が特に大きな影響を与えたのかについては不明であり、この点は今後の課題といえる。

また、各種手法による改善をしても、名詞の連続を含むアクセント句の約 12% では、推定結果に許容できない誤りがあり、改善が必要である。複合名詞では、文脈や係り受け関係によって結合の仕方に違いが出ることがあり、揺れも多い。例えば、「政権交代」が「セーケンコータイ」「セーケンコータイ」のいずれにもなり得るが、「配達証明」(ハイタツジョーメー)、 「合併失敗」(ガッペーシッパイ)のように一方しか許容されないものも多い。ここに掲げた複合名詞はいずれも単独発声では無核となる名詞で構成されているにもかかわらず結果が異なるため、これらを的確に判別するのは難しい。また、「ソ連邦解体後」「日米交換船」のようなアクセント句では係り受けの影響を受ける可能性が高い。このような問題は規則に基づくアクセント結合処理でも解決できていない問題であり、新たな手法が必要になると考えられる。

## 第6章 結論

### 6.1 本研究のまとめ

本研究では、テキスト音声合成で必要なアクセント結合処理を、可能な限り高い精度の実現を目指し、規則処理の改善としてのアプローチと統計的処理の導入によるアプローチの両面からの工夫をした。

現在のテキスト音声合成で用いられているアクセント結合規則はいまだ不備な点を含むものであり、第3章ではこの規則の改善をおこなった。まず、規則に基づいて行なわれる処理が持つ意味を的確に捉えることから、複数の規則の共通点を見つけ、同等の価値を持つ単純な規則に整理することができた。そして、整理された規則に基づいて、用言等の活用に伴って発声するアクセントの変化を規則に導入することを考え、正答率を向上させることができた。しかし、規則によって幅広い現象を捉えることには限界があり、原理上、例外的なアクセント変化に対応することは不可能であった。また、副次アクセントなど、現状の規則では扱うことが不可能な現象も存在し、規則による処理ではこれらに対応できないことが言えた。

この問題の抜本的解決のため、統計的処理の導入を必要としたが、その前段階の作業として、第4章では、統計的処理に必須となるアクセント正解ラベルを持つテキストデータベースの作成を行なった。アクセントに関する用語の定義が明確になされていない現状であっても、将来にわたり多様な研究に有用なものとなるように工夫をし、計画の約四分の一である4,166文について完成した。今後、多くの研究において活用されていくよう期待している。

そして、第5章では、作成したアクセントデータベースを用いて、条件付確率場(CRF)による学習・推定を行ない、結果として、規則による処理を上回る高い精度を実現することができ、規則から得られる知識を統計的手法に融合させた場合には特に高い精度を実現することができた。学習に用いられた文が3,581文と少ないものであるにも関わらず高い精度でのアクセント処理ができたのであり、既存の手法からの抜本的改善として成功したものであると考えている。

### 6.2 今後の課題

今後の課題と考えら得る点を以下に掲げる。

#### データベースのサイズ拡大

アクセントデータベースの作成においては、日本音響学会 新聞記事読み上げ音声コーパス(JNAS) [19] で使用されているすべての文に対してラベリングすることを予定している。現状で完成しているのはその四分の一程度にとどまり、今後も同様の作業を続けて完成させ



ることを予定している。使用できるデータの総量が増加することにより、統計的学習・推定の精度が向上することも期待できる。

### 統計的学習の改善

統計的学習によって高い精度でのアクセント処理を実現することができたが、アクセント処理に失敗するものはいまだ残り、特に名詞の連続などでは手法に不十分な点があると言わざるを得ない。アクセント処理の結果に影響を与える要素を的確に学習できるような工夫が望まれる。

### 統計的学習の結果に基づく規則の改善

条件付確率場を使った学習の結果は、用意した素性の重要度として表現される。したがって、規則に直結するように適切な素性が用意されていれば、重要度の絶対値が大きくなった素性を見ることで、改めて規則でアクセント処理を表現できることになる。このように、規則による処理と統計的処理の双方の成果・利点を互いに取り入れ、さらに必要に応じて併用することで、全体の処理精度が高められるものと考えられる。

## 謝辞

本研究を進めるにあたって多くのご指導をくださいました峯松信明助教授、広瀬啓吉教授に深く感謝致します。

アクセントデータベースの作成（第4章）においては、千葉大学の伝康晴助教授には大変お世話になりました。また、文科省科研費特定領域「日本語コーパス」の電子化辞書班の方々および音声対話技術コンソーシアム（ISTC）の方々には研究に関する助言をいただき、感謝いたしております。

その他、研究室の先輩方や同期・後輩の皆さんには暖かく支えていただき、助言も頂きました。ありがとうございました。

## 参考文献

- [1] 佐藤, 匂坂, 小暮, 嵯峨山: “日本語テキストからの音声合成”, NTT 研究実用化報告 第 32 巻第 11 号, pp.2243-2252 (1983)
- [2] 匂坂, 佐藤: “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.7, pp.847-856 (1983)
- [3] 匂坂, 佐藤: “テキストからの音声合成を目的とした日本語アクセント規則”, NTT 研究実用化報告 第 32 巻第 11 号, pp.2253-2265 (1983)
- [4] 嵯峨山 他: “擬人化音声対話エージェント基本ソフトウェアの開発プロジェクト報告”, 情報処理学会研究報告, 2003, 124, 2003-SLP-49-56, pp.319-324 (2003)
- [5] Galatea ウェブサイト, <http://hil.t.u-tokyo.ac.jp/~galatea/>
- [6] 喜多, 峯松, 広瀬: “日本語テキスト音声合成を目的としたアクセント結合規則の構築と改良”, 電子情報通信学会 技術研究報告, SP2002-26, pp.13-18 (2002)
- [7] 匂坂, 佐藤: “付属語連鎖における副次アクセントの分析”, 日本音響学会音声研究会資料, S83-05, pp.31-37 (1983)
- [8] Pierrehumbert, Beckman: “Japanese Tone Structure (Linguistic Inquiry Monograph, Vol 15)”, MIT Press (1988)
- [9] 宮崎, 大山: “日本語音声出力システムの言語処理”, NTT 研究実用化報告 第 35 巻第 2 号, pp.157-167 (1986)
- [10] 伝, 宇津呂, 山田, 浅原, 松本: “話し言葉研究に適した電子化辞書の設計”, 第 2 回「話し言葉の科学と工学」ワークショップ講演予稿集 pp.36-46(2002)
- [11] 伝, 山田, 宇津呂: “UniDic version 1.1.0 ユーザーズマニュアル” (配布元: Galatea ウェブサイト) (2003)
- [12] 日本放送協会: “日本語アクセント辞典 改訂新版”, 日本放送出版協会 (1985)
- [13] 川上 蓁: “日本語アクセント論集”, 汲古書院 (1995)
- [14] 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 (1989)
- [15] 窪園 晴夫: “音声学・音韻論”, くろしお出版 (1998)
- [16] 窪園, 太田: “音韻構造とアクセント”, 研究社出版 (1998)

- [17] 田中 宣廣: “付属語アクセントからみた日本語アクセントの構造”, おうふう (2005)
- [18] 独立行政法人 国立国語研究所: “国立国語研究所報告 124 日本語話し言葉コーパスの構築法” (2006)
- [19] 板橋 他: “日本音響学会新聞記事読み上げ音声コーパスの構築”, 日本音響学会 1997 年秋季講演論文集 I, pp.187-188 (1997) 電子情報通信学会 技術研究報告, SP2002-26, pp.13-18, 2002.
- [20] 長野, 森, 西村: “N-gram モデルを用いた音声合成のための読みおよびアクセントの同時推定”, 情報処理学会論文誌, Vol.47 No.6 pp.1793-1801 (2006)
- [21] J. Lafferty, A. McCallum, and F.Pereira: “Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data”, Proceedings of the 18th International Conference on Machine Learning, pp.282-289, 2001.
- [22] 鹿島, 坪井, 工藤: “言語処理における識別モデルの発展 — HMM から CRF まで —”, 言語処理学会第 12 回年次大会 チュートリアル資料 pp.1-10 (2006)
- [23] 工藤 拓: “CRF++: Yet Another CRF toolkit”  
<http://chasen.org/~taku/software/CRF++/>

## 発表文献

- [1] 黒岩, 峯松, 広瀬: “日本語テキスト音声合成用アクセント結合規則の改良”, 日本音響学会 2005 年秋季研究発表会講演論文集 1-Q-27 pp.427-428 (2005)
- [2] 黒岩, 峯松, 広瀬: “活用語尾に着眼した日本語アクセント結合規則の整理と高精度化”, 言語処理学会第 12 回年次大会発表論文集 P9-9 pp.995-998 (2006)
- [3] 黒岩, 峯松, 伝, 広瀬: “特定話者による大規模アクセントラベリングとそのデータベース化”, 日本音響学会 2007 年春季研究発表会講演論文集 1-Q-18 (2007) — 発表予定
- [4] 黒岩, 峯松, 広瀬: “日本語音声合成を目的としたアクセント処理のための規則と統計的学習”, 日本音響学会 2007 年春季研究発表会講演論文集 1-Q-19 (2007) — 発表予定
- [5] 黒岩, 峯松, 伝, 広瀬: “大規模アクセントラベリングコーパスの構築とそれに基づくハイブリッド型アクセント結合処理”, 言語処理学会第 13 回年次大会発表 PD2-6 (2007) — 発表予定
- [6] 黒岩, 峯松, 伝, 広瀬: “単独ラベラによる大規模アクセントデータベースの構築およびそれを利用した統計的アクセント結合処理の検討”, 電子情報通信学会 技術研究報告 (音声研 2007 年 3 月研究会) (2007) — 発表予定

## 付 録 A ラベラに配布した説明文書

第 4 章のアクセントデータベース作成に際し、ラベラに対してを示した文書を掲載する。

# アクセントデータベース作成作業について

東京大学大学院情報理工学系研究科 広瀬・峯松研 修士課程2年 黒岩 龍  
2006年9月

## 1 はじめに

アクセントデータベース作成作業にご協力いただき、どうもありがとうございます！  
この文書には、作業に必要な基礎知識、作業方法や注意点について記してありますので、隅々まで良くお読みください。ご不明な点については気軽にお尋ねください。

## 2 日本語東京方言のアクセント

### 2.1 日本語のアクセントの特徴

今回は“アクセント”についての作業にご協力いただくのですが、ひとことで“アクセント”と言っても言語によって大きな違いがありますので、まず少しご説明しておきましょう。

アクセントとは、雑に言ってしまうと、言葉を声で発するとき短い区間で何らかのメリハリを付けることを表しますが、その実際の表現形態としては強勢アクセント（ストレスアクセント）・高低アクセント（ピッチアクセント）の2種類があります。

強勢アクセントとは音の強さの変化で表現するアクセントで、英語などで使われています<sup>1</sup>。例えば、“Eng-lish”と言う際に *Eng* を *lish* より強く発音する、といったものです。

それに対し、高低アクセントは音の高さの変化によるアクセントで、日本語のアクセントはこちらです。例えば、「内野手（ナイヤシュ）」と言う際に“イ”“ヤ”の部分が高く発音する、というようなものです（東京のアクセントの場合）。強さについてはほとんど変化していないこともお分かりになるかと思います。

今回の作業で扱うのは日本語ですから、音の高さの変化についてのラベリングをしていただくこととなります。

### 2.2 東京アクセントの特徴

すでに述べたように日本語は高低アクセントですが、方言によって大きな違いを持っています。今回扱うのは東京のアクセントですので、それについて説明します。

<sup>1</sup>強さの変化に伴って高さや継続長にも自ずと変化が現れます。このことについて、英語などのアクセントは強さのみではなく複数の音響量によって表現されているのだとする意見もあります。

東京のアクセントは、モーラ（拍ともいう）<sup>2</sup>ごとの高低変化として表現されます。図1は、単語を個別に発声した場合の具体例です。

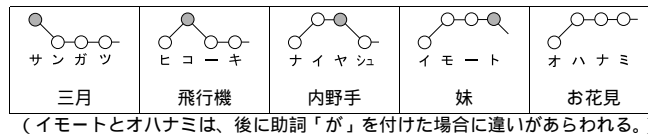


図1: アクセント型の種類（4モーラ語の場合）

図1に示された各単語のアクセントの違いは、高さが下がる（低くなる）位置の違いのみによって区別することができます。このように、音の高低変化のうち下降のみがアクセントを区別する上で重要な役割を持ち、この下降部分の直前のモーラを「（アクセント）核」といいます。それに対し、高さの上昇については、図1を見て分かる通り、1モーラ目が核である場合以外において1モーラ目から2モーラ目にかけて音の高さが上がる<sup>3</sup>ということのみであり、アクセントに関して重要な役割を果たしていません。図1は単語を個別に発声した場合ですが、文を読む場合でも類似したことが言えます。文の場合は、文の中に幾つかのアクセント的なまとまり（アクセント句）ができ、その中で前述のようなことが言えます。アクセント句の中の核の数は1つの場合が多いですが、核が無いアクセント句や、核が複数あるアクセント句（句の先頭で音が高くなった後に複数回音が下がる）もあります。

また、実際の発声においては文全体にわたる緩やかな変化（おおむね、徐々に音の高さが下がっていく）がありますが、それはイントネーションとして区別されます。疑問文などの終わりで音の高さが上がるのも、アクセントではなくイントネーションです。

### 3 ラベリングのしかた

今回のラベリング作業では、提示する文（または単語等）に、アクセント句境界位置（“/”印）とアクセント核位置（“ ”印）を付け加えていただきます。厳密な正解があるわけではありませんので、自分が読み上げる場合を想定してラベリングを行なってください。

#### 3.1 アクセント句境界位置

前述の通り、アクセント句の先頭では音の高さが上がります。具体的には、アクセント句の1モーラ目に核がある場合は1モーラ目が高くなり、それ以外の場合は1モーラ目か

<sup>2</sup>モーラとは、およそカナ1文字に相当しますが“シャ”“チャ”などは2文字合わせて1モーラとなります。“ッ”や伸ばす音の“ー”は単独で1モーラになります。

<sup>3</sup>先頭部分が長音などの場合（例：紅茶）では、上昇がほとんど見られずに1モーラ目から高い音になる場合があります。



ら2モーラ目にかけて音の高さが上がります。このような変化を目安にアクセント句境界の位置を見つけ、“/”印を付けてください。休止が入った場合も、通常はアクセント句の境界が生じます。アクセント句の境界の発生は読む速さの影響を受けるため、こちらの提示する話速目安(1モーラあたりおよそ150・200ミリ秒)に合わせて自然に読んだ場合を想定して印を付与してください。

「私は踊った」の場合...

- ワタシワノオドッタ とすると、“オ”から“ド”にかけて上昇があることを示す。(その準備として“オ”が前の“ワ”より低くなっている場合も多い)
- ワタシワオドッタ とすると、“ワ”“オ”“ド”のあたりに上昇も下降もなく滑らかにつながっていることを示す。

「私は食べた」の場合...

- ワタシワノタ 'ベタ とすると、“タ”で(若干の)上昇があることを示す。
- ワタシワタ 'ベタ とすると、“タ”は前の“ワ”から上昇せず滑らかにつながっていることを示す。

「彼は踊った」の場合...

- カ 'レワノオドッタ とすると、“オ”から“ド”にかけて上昇があることを示す。
- カ 'レワオドッタ とすると、“ワ”“オ”“ド”のあたりに上昇も下降もなく滑らかにつながっていることを示す。

「彼は食べた」の場合...

- カ 'レワノタ 'ベタ とすると、“タ”で上昇があることを示す。
- カ 'レワタ 'ベタ とすると、“タ”は前の“ワ”から(低い音のまま)上昇せず滑らかにつながっていることを示す。

ここまで述べた例では、いずれも、どちらになるかは読む速さなどにより異なります。読む速さをおよそ目安に合わせて自然に読んだ場合を想定し、あまり考え込まずに“/”印を付けるようにつけてください。

### 3.2 アクセント核位置

前述のとおり音の高さが下がる箇所の直前のモーラがアクセント核です。作業としては、アクセント核の直後、つまり音の高さが下がる位置に“'”印を付けてください。

アクセント核の位置が分かりづらい場合、まず、該当するアクセント句を自然に滑らかに読む様子を想定してアクセント核位置をおよそ見込みます。そして、改めて、アクセント核と見込んだ箇所は急激に音の高さを下げそれ以外は同じ音の高さで平坦に(イント

ネーションを除去して)読むようにし、違和感が生じるかを確かめます。違和感が生じないようなアクセント核位置が定まったら、その位置がアクセント核ですので、“ ” 印を付けてください。

アクセント句内に複数のアクセント核を見込んだ場合にも、アクセント核と見込んだすべての箇所可能な範囲でできるだけ急激に音の高さを下げて同様にします。この際、イントネーションによる音の高さの下降を誤ってアクセント核としてしまわないよう、注意してください。例えば、「内閣」は“ナ 'イカク”であり、“ナ 'イ 'カク”や“ナ 'イ 'カ 'ク”ではありません。

### 3.3 注意点

- 疑問文の終わりや、文末の「か」「ね」などは音の高さを上げずに読んでください。同様にアクセント句末を上げるような読み方(「あなたが」の“が”を上げるなど)をしないでください。
- 母音が無声音にならないようにしてください。「食べます」の“す”は無声音(sのみ)にすると音の高さが下がらないように感じますが、きちんと母音まで発音する場合を想定して“タベマ 'ス”としてください。
- 文字通り読んでください。「フィルム」(「フィルム」)のようにア、イ、ウ、エ、オ、ヤ、ユ、ヨが付いたものは1モーラになるように読まなくてはなりません。同様に、伸ばし棒(ー)などの無いところで勝手に伸ばしてはいけません。
- 促音(ッ)のところにアクセント核がある場合は、ッの前に“ ” 印を付けてください。また、伸びる音(“ー”の付いた音や“ネエ”などのように同じ母音が続く場合)で伸ばしながら音の高さが下がる場合も、同様に“ー”などの前に“ ” 印を付けてください。ただし、同じ母音が続いて、「用いる」などでは長音とはならないために“モチ 'イル”でなく“モチイ 'ル”となります。

## 4 作成していただいたデータの取り扱いについて

この作業によって作成していただいたデータは、研究用として公開する予定です。あらかじめ、ご承知おきください。