

日本語学

1998.MARCH
VOL.17

3

明治書院

特集 アクセント研究の現在

話しことば、音声言語の教育が重要性を増してきた現在、
アクセント研究の現状を探る。

◆連載

続・落語のレトリック

ユーミンの言語学

気づかない方言

文法を考える

話しことば

—その研究と教育を—



アクセント
研究の現在

イントネーションと文法…おなじ親の子

ニック・キャンベル

1 はじめに

プロソディあるいは韻律はイントネーションとタイミングを含み、機能的意味を示す。音声言語学では、イントネーションは声の抑揚を示す情報とされ、高低二つのトーンの組み合わせでことばや意味のバリエーションを表現すると考えられている(注1・2・3・4)。従来、プロソディと文法の関係においては、プロソディは文法に依存するもの、すなわち下位構造ととらえられ、このような観点からの研究は多く進められてきた(注5・6)。

テキスト文では意味を伝えるために、文法的構造つまり統語論と意味論が大きな役割を果たす。朗読では、韻律

は文法に従属する。しかし自由発話の場合、特に対話では韻律は独立的に働き、両者の役割は変わる。

自動発話処理の観点から見れば、プロソディ(以下韻律と文法の関係は、むしろ並立的関係ととらえられる。本稿では、イントネーションは主に基本周波数(音高)から、タイミングは音韻時間長から韻律的特徴を抽出し分析した結果を用いて、文構造と意味と韻律の関係について述べる。

2 曖昧性の解釈

同一単語列であっても、韻律の違いによって別の意味

を伝えることができる。例えば、「あさ」は「んをたべる」という文字列は、二つの場面(朝に/朝ご飯)が考えられる。「あさ」の後にポーズがあれば、もちろん意味の区別はしやすいが、ポーズがなくても意味を区別することができる。この場合、声の高さ、つまりイントネーションだけではなく、タイミングも含めて違いを表す。

意味的情報を持つ韻律とは、端的にはイントネーションとタイミングの組み合わせである。したがって、ここでは韻律を二次元的構造、つまり縦軸をイントネーション、横軸をタイミングと考え、韻律と文法の関係を明らかにする。また、イントネーションはセリアンス(顕在性, salience)を示し、タイミングはセグメンテーション(分節, segmentation)を示すことを提案する。セリアンスは意味的な比重、セグメンテーションは意味単位の区切りであり、二つの要素によって意味を伝えると考えられる。

[1] セリアンス

文法的単位および韻律的単位は構造的に一致する場合が多い。このため韻律は文法に依存的であるように見えるが、文法の影響のみでは説明できない場合もある。

例えば、「兄に電話したい。」という発話は、「妹ではなく、兄に(a、実線)」あるいは「手紙ではなく、電話で(b、

破線)」という二つの意味が考えられる。前者(a)は「兄」に強調があり、基本周波数を見ると(図1)、三つのピークのうち、初めが最も高いが、後者(b)は二つ目のピークが最も高くなる。ここでは同じ文法的単位が、発話意図によって、異なる韻律的単位を示す。すなわち、声の高さの変化(イントネーション)によって、強調的意味の違いを表現する。これが韻律の縦軸としてとらえるセリアンスである。

[2] セグメンテーション

発話された文の意味を正しく理解するためには、文の構成単位ごとに適切に区切らなければ、その意味内容まで変わってしまう。例えば、「古い本と雑誌」では、「古い」という形容詞がどの名詞に係るかあいまいである。しかし、「古い本」と「雑誌」、あるいは「古い(本と雑誌)」と韻律区切りをつけることによって、意味の違いを示すことができる。

これらの区別は、イントネーションだけではなく、タイミングつまりは音韻時間長の伸びもかわる。図2は「古い(本と雑誌)」に相当する例であり、矢印は韻律境界を示す。音声波形には二箇所(句末伸び)が見られる。「本」と「雑誌」との間の基本周波数に見る区切りは音声波形には

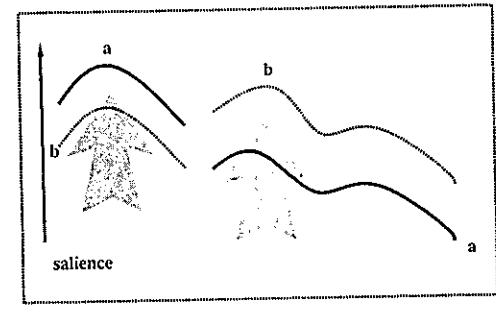
ないため「本」と「雑誌」は一つの意味単位となっている。これが韻律の横軸としてとらえるセグメンテーションである。「古い」の後にポーズがなくても、母音/i/の伸びによって、韻律境界を示す。このように、韻律区切りは意味単位を表し、時間軸に沿って働く(図3)。

人間はこれらの韻律的特徴から意識せずに意味の違いを識別する。また逆に、意味の違いを表すために、イントネーションやタイミングを使う。機械がこのような韻律情報を処理するためには、音声波形からどのような音響的特徴を抽出する必要があるだろうか？

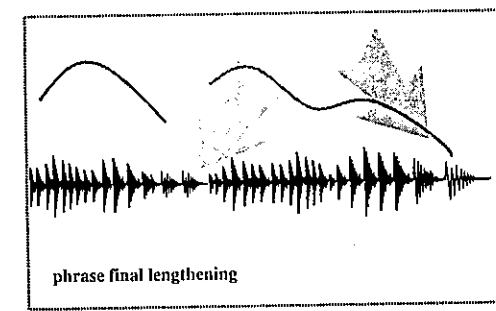
3 韻律情報からの意味抽出

前述のような意味の違いを区別するために機械処理の場合、例えば音声翻訳システムにおいては、入力音声の韻律を抽出することによって強調の位置を明らかにし、意味を区別することができる。

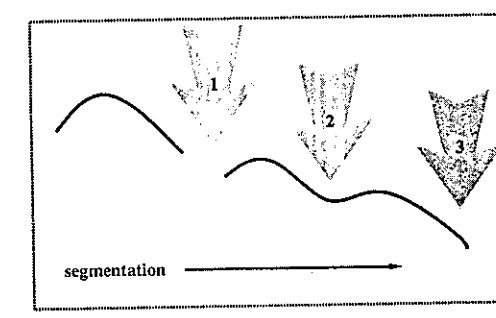
セイリアンスの具体的な例として、自動強調抽出を取り上げる。この手法は人間の強調を含んだ発話音声サンプルと音声合成の強調を含まない発話音声サンプルの基本周波数を比較し、強調箇所を抽出した筈¹⁾。使用サンプルデータは、一名の男性話者によって朗読された会議



〈図1〉イントネーション、基本周波数はセイリアンス及び強調を示す。



〈図2〉音声波形に見る句末伸び、波形が二ヶ所で水平に伸びた状態になっている。



〈図3〉タイミング、セグメンテーションあるいは韻律区切り。ここで三つの韻律境界が見える。

登録の対話文の八九文である。さらに、コントロールとして、人間の強調を含まない自然発話音声も使った。この対話文は以下のようにすべて同一の文字列から構成されており、強調の位置だけが異なるものである。

- 1…今回は 割引を おこなっております。
- 2…今回は 割引を おこなっておりません。
- 3…今回は 割引を おこなっておりません。

まず、強調を含まない音声(右記1)と強調を含んだ音声のそれぞれに対して、基本周波数曲線から各音韻ごとの平均基本周波数を求め、それぞれの音韻の平均基本周波数からの移動量を求めた。次にその移動量の差を求め、差が最も大きい音韻を含む文節に強調があると判断した。その結果、自然音声対自然音声では八一・三%(六四分の五)の正解率となり、合成音声対自然音声では七五%(六四分の四)の正解率を得た。このように高い精度で強調箇所が抽出されたことから、基本周波数の情報のみから発話の意味の違いを識別できることが分かる。したがって、韻律の縦軸は意味的な比重を表し、基本周波数でセイリアンスを示すと言えよう。

[1]タイミングと音韻時間長
セグメンテーションの具体例に入る前にタイミングに

ついて説明する必要がある。繰り返しになるが、従来の研究では基本周波数が韻律の中心課題となり、タイミングを取り上げた研究は多いとは言えない。その理由は分析データが少ないことに重ねて、生データからタイミングのパターンを見ることが困難であるためと言える。

[1] 音韻時間長の個性

〈表1〉日本語に於ける子音と母音の平均時間長と標準偏差

consonant	mean	S. D.	n
/s/	89.05	18.55	521
/t/	75.32	51.91	1187
/h/	66.71	21.19	334
/k/	60.04	28.41	1152
/y/	56.24	14.61	313
/m/	45.98	11.35	763
/w/	42.84	11.39	303
/n/	41.38	12.24	1119
/r/	23.98	10.12	1105
vowel	mean	S. D.	n
/N/	90.99	24.38	450
/a/	83.86	22.54	3380
/e/	80.04	21.18	1697
/o/	77.71	23.59	2359
/i/	69.82	23.53	2095
/u/	58.32	23.02	1426

「う」も同じ長さであると知覚されるが、実際は表1が示すように、母音/a/（平均時間長八三・九ミリ秒）は母音/u/（平均時間長五八・三ミリ秒）より長い。このように、音韻毎に平均時間長を比較すれば、かなり大きな差が認められる。

この差は音韻を発声する際の顎の動きで説明できる。あごや舌の動きが必要になるほど、発声に要する時間は長くなる。例えば、母音/a/は、母音/u/よりも発声する際に口を大きく開けるため、平均的に音韻時間長は長い。また、子音/s/は子音/r/よりも、口蓋と舌先との間が適切な位置で摩擦音を発声しなければ認識されないため、長く時間がかかるとされている。

このように音韻には個別の特徴がある。音韻ごとの長さだけを見ても伸びは分らない。同じ七〇ミリ秒の音韻時間長であっても/a/については短い、/u/については長い伸びとなる。つまり音韻の伸びを比較する場合、音韻の種類による個性を除かなければならない。

[1]2 音韻時間長の正規化

音声波形から計測した物理的な長さ（絶対値）を用いて、タイミングを比較するためには正規化する必要がある。正規化とは、音韻固有の時間長を排除し、平均的な偏差を

算出することである。

音韻の種類ごとに平均と標準偏差を算出して、それぞれの音韻の正規化した継続時間長伸縮尺度値（Z値）を決定した。日本語の主な母音と子音の正規化前の値と正規化後の時間長と伸びを図4と図5に示す。左図は実測時間長、右図はZ値である。

平均値より標準偏差分だけ高い値を+1に、低い値を-1と表したとき、ほとんどの音韻が-3から+3の範囲に納まる（音韻時間長の分布が正規分布であるとすると全体の九九・七%）。音声波形から計測した値を正規化することにより、音韻固有の時間長特性に依存しない各音韻の時間伸長つまり「伸び」を見ることができるとなる。

発声の生理的影響による各音韻毎の継続時間長の差は正規化することによって、平均値や分散の異なる音韻間の比較が可能になる。

[2]1 モーラタイミングと「伸び」

セグメンテーションの具体例は、モーラタイミングの検討を素材とする。このモーラ枠を基準にリズムや伸びのパターンが見えてくる。

日本語は一般に、モーラタイミング（V・CV・CCVなど）から構成されるユニット）もしくは音節タイミング言語と

言われている（注8）。モーラ枠内の正規化された音韻時間長の伸びを見ることによって、タイミングによる影響を明らかにする目的で大規模な日本語音声データベースを用いて音韻時間長の分析を行った（注9）。正規化した時間長伸縮尺度（Z値）を用い、隣接音韻環境に基づく子音、母音の平均伸縮尺度値の分散分析を行う。その結果、モーラを構成する母音は、子音の音韻長が短い場合には伸長し、逆に長い場合には短縮することによって、モーラ枠を均一に保つ傾向が見られた。

例えば、/s/や/t/のような音韻長の長い子音に後続する母音はモーラ枠の影響で短くなる。時間的に短い子音に後続する母音は伸びることが多い。すなわち、モーラ単位を均一なタイミングにしようとする補償的作用が働く。図6は正規化された音韻時間長でこの動きを示す。図6の右図については以下の文末伸びで説明する。

[2]1 韻律境界の句末伸びについて

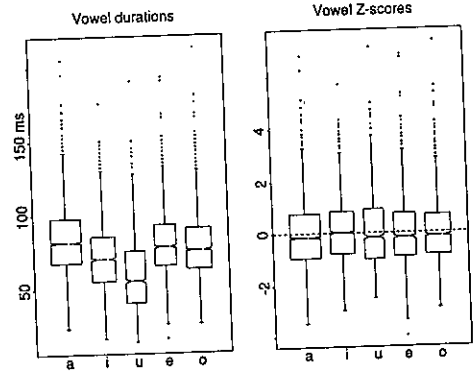
次に、セグメンテーションのもう一つの具体例としてモーラ枠の影響による句末伸びを見る。注10において述べている助詞の「は」(wa)の考察から日本語の句末伸びを例示する。なお、分析データは日本の雑誌および新聞から音韻バランスを考慮して抽出されたプロアナウン

イントネーションと文法：おなじ親の子

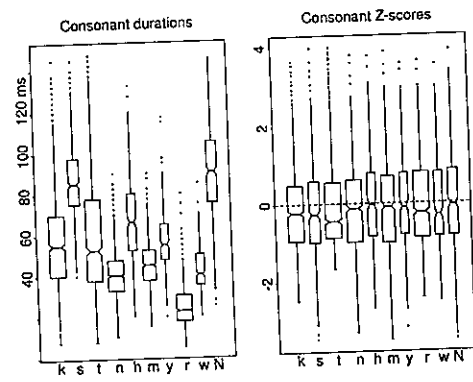
サーによって発声された五〇三文音声からなる(注1)。
 日本語においては、助詞は句末に、動詞は文末にあることを前提に、モーラ枠の影響による句末伸びならびに文末伸びを見るため、すべての音韻のタイミング特徴を品詞グループごとと比較した。このような比較は正規化以前の音韻サンプルでは、個別の母音や子音の差が大きいため不可能である。音韻時間長を正規化することにより、異なる音韻を一つのグループにまとめて分析する。まず始めにこの個別差を除けばグループ別の詳細な比較が可能になる。

図7の上図左はすべての/a/と/w/の音韻時間長の分布を示す。実測値における/a/の時間長は/w/より長い。しかし、正規化した後(図7右)はすべての/a/と/w/の音韻の種類による個別差がなくなり、正規化した平均と標準偏差は同じとなる。図7の下図は「助詞」のグループ別分布を示し、左図は実測値の音韻時間長、右図はZ値である。ここで母音/a/の句末伸びは明らかであるが、/w/に伸びが見られない。一見、句末伸びは「句末母音のみ」と思われるが、必ずしもそうとは限らない。

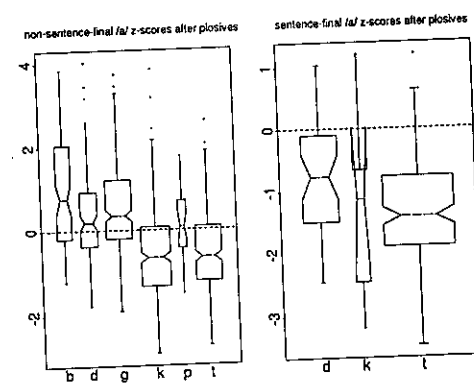
分析データを含む三〇三の/w/のうちの二二七(七割以上)は助詞の「は」に含まれている。品詞的分布から見



〈図4〉五母音の平均時間長とその正規化した結果：ボックスにおける水平線は25% 50% 75% パーセンタイルを示す。ボックスの幅はサンプルにおけるトークン数に対応する。



〈図5〉子音の平均時間長とその正規化した結果



〈図6〉破裂音の後の/a/の伸び：有声破裂音が平均的に短く、/a/が伸びる。無声破裂音が平均的に長く、同じモーラの/a/が縮む。文末の分散の異なる音韻間の比較が可能になる。

れば、ほとんどの/w a/音(「わ」と「は」は助詞の「は」であるため、平均から伸びの比較は容易ではない。句末伸びが多いため、伸びている方が普通になる。すなわち、句末に現れる助詞の「は」のモーラ単位は全体的に伸びがあると考えられる。

[2] 2 文末伸びについて
 さて、動詞に表される文末の伸びはどうなるだろうか。図8の左図が示すように実測値で文末母音を見れば、どれも大差はないが、上記のように母音/a/は普通に母音/i/と/u/より長い。Z値で見れば、文末の/a/はかなり短くなっている、この母音/a/が伸びていないことをモーラ単位の先行子音の存在で説明することができる。

前述したように日本語はモーラ言語であるため、音韻でなくモーラを単位として検討する必要がある。そのため、母音だけでなく、先行子音も分析する必要がある。図9は文末モーラにおける/a/の先行子音だ、か、たの伸びを表す。ボックスの幅はサンプルにおけるトークン数を示すように、「た」の/t/が多く、すべてに大きな伸びが認められる。

上記のように、長い子音はモーラ枠の影響により短く

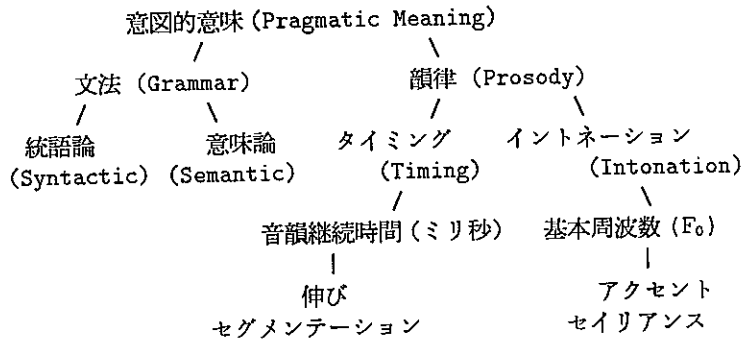
イントネーションと文法：おなじ親の子

音声言語が持つ意味における機能的体系を、図10に示す。従来イントネーションは文法に依存していると考えられてきたが、ここでは図10が示すように文法と韻律を同レベルとして位置付けたい。文法における統語論と意味論の関係は、韻律におけるタイミングとイントネーションと対置される。また、統語論はセグメンテーションと関連し、意味単位を決定する。意味論はセイリアンスと関連し、意味的なウエイトの関係を決定する。つまり、韻律は文法と兄弟のような関係にあり、その親は意図的な意味である。

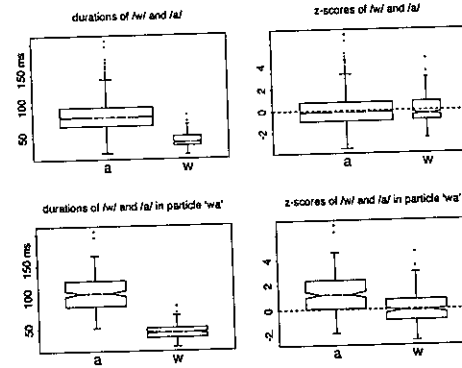
韻律を実現するためには、意味的な情報を階層的に分類する必要がある。最下位層は音韻記号で識別可能な音

4 イントネーションの役割

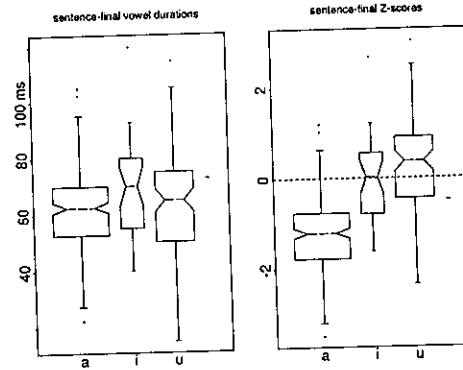
このように、タイミング情報の分析によって、モーラの影響を考察し、モーラ枠と文末伸びを例として文構造とタイミングの関係を明らかにした。つまり、タイミングはイントネーション同様、意味構造の表す要素である。



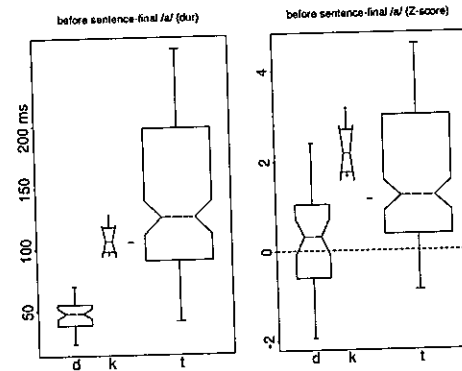
〈図10〉 韻律構造と文法構造との関係



〈図7〉 母音/a/と子音/w/の伸び、左は音韻時間長(ミリ秒)、右はZ値。上は全てのサンプル、下は助詞のみである。



〈図8〉 文末母音の音韻時間長と伸び



〈図9〉 文末モーラに於ける子音

韻特徴レベルの層であり、その上は同音語をアクセントで識別するアクセントレベル層である。この二つの層は辞書的レベルである。

次に続く層は今回の検討の中心であり、文中の単語列を韻律によって区別する韻律区切りおよび強調のレベルである。同一単語列からなり、係り受け関係があいまいな文の区別をする。

さらには、発話行為レベル、話者感情レベル、信憑性や信頼性などを識別するレベルなどが続く。(詳細については注12を参照されたい)

5 むすび

本稿で述べた分析手法は、意味の違いを識別する方法として有効であることは既に確かめられている。人間は、韻律が持つ意味を無意識のうちに理解することができ、が、音声翻訳システムなどに入力音声を理解させるためには、韻律情報を層的に分類するためのアルゴリズムが必要である。またここで利用した正規化は、人間が行う知覚的な認識の方法というよりも、計算機で扱うための分析手法である。しかし、人間の音声認知においても似たような正規化を行っていると言えないだろうか。

〔謝辞〕本稿の執筆にあたり、馬塚れい子さん、佐藤寿子さん、特に早川久美子さんに御協力いただいた。この場を借りてお礼を申し上げます。

注

- 1 杉藤美代子(一九八二)「日本語アクセントの研究」(三省堂)
- 2 Haraguchi, S.(1977) "The tone pattern of Japanese: an autosegmental theory of tonology" Kaitakusha, Tokyo.
- 3 Kubozono, H.(1993) "The Organisation of Japanese Prosody" (くま)お出版
- 4 ニック・キャンベル(一九九七)「Tobiシステムと日本語への適用」(日本音響学会誌、五三巻三号)
- 5 樋口宜男(一九八二)「日本語連続音声における単音の持続時間に関する研究」(東京大学、博士論文)
- 6 句坂芳典(一九八五)「音声合成のための韻律制御の研究」(早稲田大学、博士論文)
- 7 北川・キャンベル(一九九八)「文音声における波形比較による強調抽出」(日本音響学会講演論文集)
- 8 Rubetzkoi, N., Principles of Phonology, 1969.
- 9 ニック・キャンベル・句坂芳典(一九九二)「音声タイミングにみられるモーラと音節の影響について」(音声言語研究会資料、信学技報SP九〇〜一〇七頁)
- 10 Campbell, W.N.(1991) "Segmental Elasticity and Timing in Japanese Speech." Speech Perception, Production and Linguistic Struc-

- 11 阿部匡伸、句坂芳典、梅田哲夫、桑原尚夫。(一九九〇)「研究用日本語音声データベース利用解説書」(連続音声データベース)。TR-0166, ATR 自動翻訳通信研究所
- 12 ニック・キャンベル(一九九七)「プラグマティックイントネーション」『文法と音声』(くま)お出版
(Nick Campbell ATR 音声翻訳通信研究所)

読者の皆へ

● 投稿のご案内
本誌へのご意見・ご要望、街でひろったことばに関する話題、日本語についての疑問など、左記「日本語学」投稿係までお寄せください。

- ・「読者のコーナー」(日本語Q&A)に掲載させていただきます。
- ・原稿は四〇〇字以内でお願いします。
- ・採用分には薄謝進呈いたします。
- ・原稿はお返しいたしません。
- ・住所、氏名、職業(勤務校)をご明記ください。

● 論文投稿規定

- ・内容：日本語学に関する論文とします。
- ・分量：八〇〇字程度。
- ・採否：原稿は、本誌編集委員の審査により、採否を決定します。ただし、審査内容は一切公表しません。
- ・原稿は返却いたしません。
- ・掲載の場合、掲載誌五冊を進呈いたします。

明治書院 『日本語学』編集部
〒101 0054 東京都千代田区神田錦町1-16