

5. 韻律解釈における基本単位

—音声文法構築へ向けて—

Nick Campbell

あらまし

本稿は、韻律特徴の記述を取り上げ、音声言語のための文法的枠組みについて述べる。具体的には、機能的意味の階層論に沿った韻律の役割から、韻律記述法を論じる。計算機は意味そのものを理解することは可能ではない。したがって、発話内容の「意味」を解釈するため、音声から得られた情報を符号化する必要がある。音声翻訳の場合、音声波形から音声に含まれる韻律を抽出し、その物理量を記号としてテキストに記述する。既にこれまで韻律記号付与法として、ToBI手法が提案されている。そこでToBIは何を示すか、あるいは示しうるかという問題を検討し、意味の符号化の方法を提案する。

1. 背景

ATR音声翻訳通信研究所で進めているヒューマン・コミュニケーションにおける音声翻訳通信技術の基礎・基盤研究に立脚し、音声翻訳システムを通して、音声言語の科学について考える。ここで扱う音声は発声された言葉、なかでも朗読音声ではなく人間と人間の自然対話を対象とした人間における音声言語コミュニケーションである。文字言語と音声言語における大きな相違を韻律の役割に求め、本論で述べる韻律とは声の抑揚のみではなく、リズム、「言い方」の違いなども含むものとする。

人間の音声によるコミュニケーションを計算機で処理する場合、情報の符号化は必然である。音声には文字言語にない様々な情報が含ま

れる。したがって音声翻訳処理では入力音声を文字列表記した後、さらにより深い意味の解釈を実現するために音声に含まれる韻律を抽出する。つまり、より克明に話された「声」を写実するのである。そしてそれに従い意味情報を解釈し、記号付与、翻訳言語の韻律生成が行われる。

近年、急速に進められているこの知識ベースの大規模化およびその整備の充実に伴い、韻律情報の自動処理は、音声翻訳の実現のみではなく広範な領域・分野での活用が考えられ、そのような状況を背景にそのためのシンタクスの決定、すなわち音声文法構築への期待も大きい。

以下、韻律と意味について述べる。意味とは複雑である。同時に韻律もまた複雑である。それに対して韻律記述の情報は単純であるが、差異に着目することによって複雑な意味の構造に迫りうる。正規化による基準パタンの決定から深い意味を解釈する方法について述べたい。

2. 韻律と意味

本章では韻律を言語処理の観点から捉え、意味との関連を考察する。

2.1 音声言語処理

はじめに音声情報処理の面から簡単に説明する。音声翻訳のための主要技術は音声認識、言語処理、入出力の両エンドに関わる韻律処理および、音声合成の4つのモジュールからなる。話された言葉つまり入力音声はまず、音声認識モジュール部によってその音声が認識され、単語列として文字表記される。しかしこの段階では「単語のみ」の情報であるため、意味解釈上、曖昧性が含まれている。例えば、自然対話では、「ビール飲みへん」といった場合、飲みませんという否定の意味ではなく、一緒に飲もうよという提案を意味することも多い。あるいは、疑問の助詞「か」を付加する代わりに声の抑揚によって意味の

違いを示す場合も少なくない。

機械は、人間が行うように発話された音声を聞いて、自然にあるいは無意識的に話者の意図や目的とする意味まで識別することはできない。それにもかかわらず、入力者である話者は同じ文表現などを使って話者意図や機能的意味を言い分ける。その際、特に統語論的情報やフォーカル・プロミネンス(プロミネンスもしくは強調)は声の抑揚などによって示される。例えば、「そうですか」「そうですね」は表現の違いによって複数の意味を表す。つまり、声の高さ・リズムを使って、疑問・あいづち・ためらい(躊躇)などの発話意図を示す。このような発話意図などによる違いを識別するために、入力時の音声に含まれる韻律情報が抽出される。例えば、韻律区切りの深さは文法的構造(枝分かれなど)などを示す。また、すべての文に該当するわけではないが、上記のような同一文字列の場合や、さらには後述のより高次の意味解釈を目指す機械処理システム構築のために、韻律処理は極めて重要な役割を担う。したがって韻律記号付与は、音声情報処理における単語列から意味列への変換過程を意味し、音声翻訳システムにおいて認識された情報を言語翻訳モジュールに渡す情報としては、テキストレベルに加えて、韻律付与記述が必要である。

2.2 韻律の機能的役割

音声翻訳の入力は現在、文単位のみである。したがって当然、文音声内の情報のみしかアクセスすることができず、対話に至るまでの経緯やその背景は考慮することができない。これは機械が扱う自然対話と人間のそれとの大きな相違である。それに対して韻律と意味との関係は、音声翻訳内に限定されるものではなく、同一言語間はもちろん、人間と人間との音声言語コミュニケーションに汎く関わって言えることである。

表1 意味層と韻律の関係(くろしお出版1997『文法と音声』より抜粋)

段階	例	機能(意味の層)	実現
L0	「はし／かわ」	segmental(音韻的)	(phonetic) 音素特徴
L1	「ほし／はし」	lexical(辞書的)	accent アクセント
L2	「年輩の男と女」	syntactic(統語論的)	chunking 区切り
L3	「田中さんの本」	semantic(意味論的)	prominence 強調
L4	「そうですね」	attitude(態度的)	speech act 行為
L5	「嬉しいわー」	emotion(感情的)	interpersonal 信頼性
L6	「すきですね」	sincerity(誠実性的)	commitment 人間関係
L7	「もしもし」	individuality(個人性)	(+voicequality) 音源特徴

まず、韻律情報を峻別するために、韻律と意味との関係とその機能的役割によって階層化する。表1にこの「プラグマティック・イントネーション」(PI理論)から捉えられる意味層を示す。ただし、韻律はあくまでもゲシュタルトなもので個別に機能するものではない。理論的区別に基づき表出される機能的意味を階層として捉え、言語的意味レベルから個人に依存する意味レベルへの移行として表した。なお、意味層およびプラグマティック・イントネーションについては、『文法と音声』^[11]において詳しく述べているので参照されたい。

3. 韻律の記述

韻律情報の分析とその記録のための処理としては、(1)それぞれの韻律パラメータの計測、(2)カテゴリごとの分類が行われる。この節では、韻律情報の記述の観点から、国際的基準化を考慮したToBIシステムを取り上げ、韻律情報処理について検討してみたい。

3.1 ToBI韻律付与手法

ToBI韻律記号は^{[2][3]}、韻律特徴を記述する方法の一つとして提唱され、音声データからの韻律情報をテキストに付与する方法である。周知のようにこのシステムは、TSMモデル^[4]の理論を基本として、韻律の国際共通記述法の確立、同意を求める専門家らによって考案された。

既にまたToBIは、英語をはじめ複数の言語に対してその有効性が確かめられており、日本語への適用についても紹介されている^[5]。なお、ToBI、G-ToBI(ドイツ語版)^[6]、J-ToBI(日本語版)^[7]等については、<http://itl.atr.co.jp/ToBI>を参照されたい。

ToBIはすべての韻律特徴を記述するものではないが、次にあげる3つの基本原則に従い、必要最小限のものを記述する^[2]。第1原則は韻律のすべてのカテゴリについて記述する。第2原則はカテゴリの記述にとどめ、量的記述は行わない。第3原則は言語情報から予測できるものは記述しない、である。

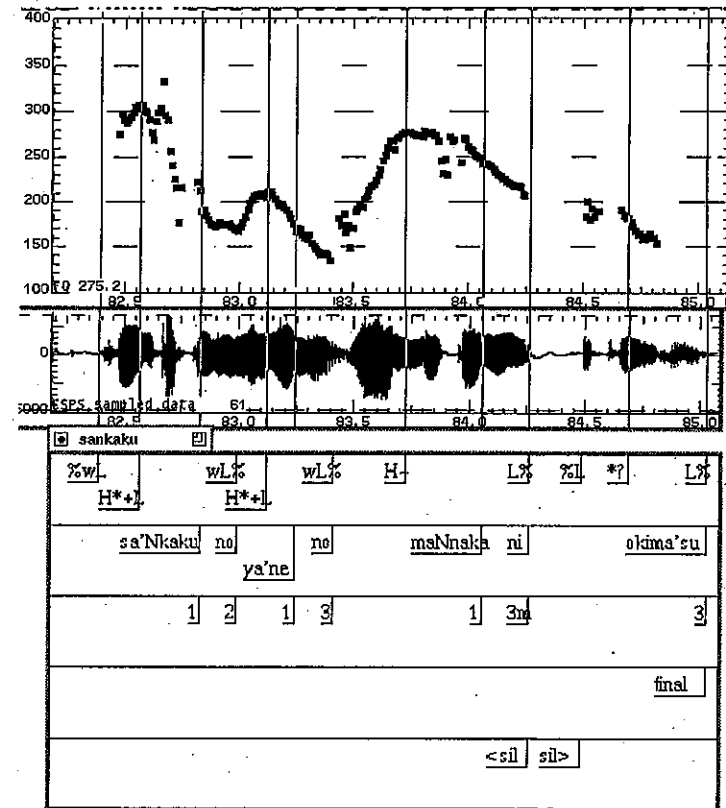


図1 ホームページからのJ-ToBIの例

上記の原則に従い、この韻律記述システムは主に2つのカテゴリで記述される。すなわち、トーンとブレイク・インデックスである。原則にも見るようにこれらは共に量的記述ではなく、韻律事象をシンボル記号によって記述する方法を採っている。図1はToBI記述の例であり、上述の2つの基本カテゴリの他に、当然単語列を示す単語層と、言い淀みなどを書き込む第4層がある。最上層はトーンの層で、アクセント、フォーカス(フォーカル・プロミネンスの略、あるいは強調という)、疑問形などを読み取る声の抑揚が示され、第3番目の層が、発話リズムやタイミングを示すブレイク・インデックスの層である。記号についての詳細は省略するが、トーンとは声の高さ(ピッチ)を表し、ブレイク・インデックスとは単語と単語の結合度を、0から4の5段階スケールで表している。なお、このような方法でToBIは、音声波形とF0曲線を基本情報として韻律を符合化し記述を行うものであるが、後述する5節の実験でも用いているように、ToBIの記号表現を利用したF0曲線を予測するアルゴリズムである。

3.2 ゲシュタルト理論の韻律

ToBIの韻律付与方式は、上記のように基本的には2つのカテゴリによる2次元記述である。つまり、声の抑揚を扱う縦軸と、リズムやタイミングを扱う時間に沿った横軸によって韻律のバリエーションを示す。従来の研究の観点からピッチに注目した成果が多く見られる。しかしピッチ以外にもタイミング等を表す重要な情報があり、それらはゲシュタルトとして複合的に関係し合う。

ToBIを用いて、どこまで韻律のバリエーションを示せるかについて、前述の韻律と意味の階層(PI理論)に従って簡略に結論的に述べておく。結論的には表1のL1~L3までは記述可能である。つまり、L1は単語アクセントレベルである。L2は構文レベルであり、チャンキング(意味単位で行われる韻律的区切りと連続の表れ)すなわち、統語論的構造を示す。ここまではToBIで充分記述できる。しかし次のL3、

すなわちフォーカスレベルは、第5節の実験を通して詳述するが、ToBIの2次元記述であるための制約が残る。それでもなお、少なくともこのL3までは、計算処理を含むことによって記述することは可能である。さらに上位のレベルについては、例えば、感情(L5)などは、今のところToBIは記述の手段を持たない。感情音声と韻律の関係に関しては現在明らかにされつつあるが、別の機会に譲ることとする。L5より上位レベルについてもなお、理論的には区別できるが自動処理の観点からも現時点では難しい。

4. 韻律処理

人間は韻律をどう解釈するかは別として、人間における韻律情報処理のメカニズムの解明ではなく、計算機における意味解釈処理について以下に述べてみたい。音声翻訳における韻律情報処理は、抽出、変換、生成の3つのプロセスからなる。抽出については第5節の実験1, 2において述べる。変換についてはここでは取り上げないが、例えば多言語翻訳の場合では、入力言語の韻律抽出結果を翻訳された単語列にマッピングして出力言語に意味を伝えることである。生成とは韻律記号を利用して意味を出力に反映させるものであり、韻律のパタンの分類化が基本である。例えば、後述の実験では検証に際し、人間の声を基準音声として用いているが、実際の場面である音声翻訳システムの場合には入力サンプルは存在しない。したがって、ToBI予測によってパタン選択を行い基準音声を決定している。

4.1 韻律のパラメータ

韻律情報処理過程の記述について述べる前にまず音声から何を測るか、韻律の主なパラメータに関して説明する必要がある。音声言語には様々な情報が含まれているが、主要なパラメータとしては、(1)声の高さ、すなわち基本周波数(通常、F0と略記する)、(2)声の大きさ、すなわちパワー、(3)話す速度、すなわち音韻継続時間長、があ

る。このうち、F0とパワーの2つのパラメータは、音声波形のみで直接計測できる。これに対して音韻継続時間長は、発話内容のテキスト化、音声認識、ラベリング処理の後に得られるもので、直接計測はできない。さらに実測値については比較のために正規化を行う必要がある。

4.2 パラメータの正規化

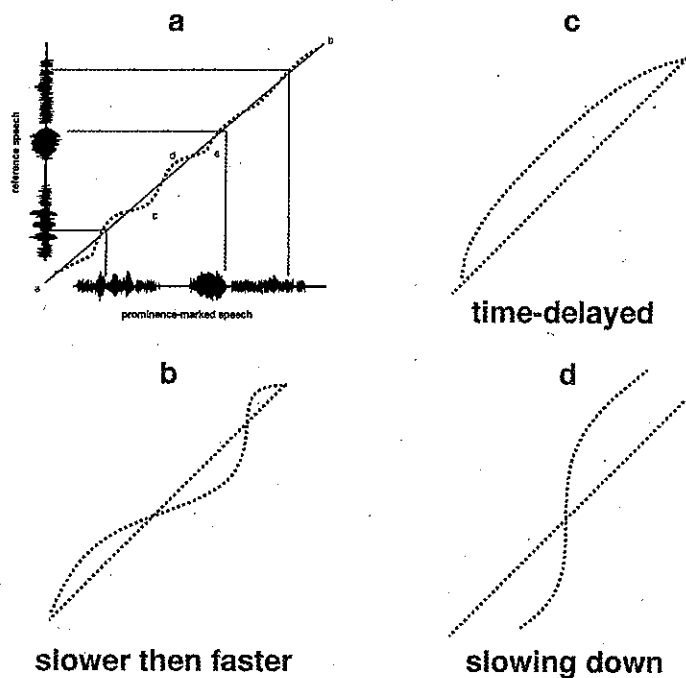


図2 (a)動的時間伸縮による2つの音声信号比較

図3 (b,c,d)DTWに表れる3種類の速度変化パターン

パラメータの正規化手法の一つに動的時間伸縮DTW (dynamic time-warping)がある。この手法は2つの音声波形を直接、比較可能にする方法で第5節の実験でも用いられているため、先に説明する。DTW手

法は具体的には2つの音声波形信号を各フレームごとに信号の特徴(ケプストラム変換)を一致させ、両者の差を数値化して比べるものである。図2と図3はDTW手法による比較の概要を示す。2つの音声波形信号は基準とする音声波形(reference speech)を縦軸に、比較対象音声波形(prominence-marked speech)を横軸にとっている。つまり、2つの音声波形信号は言い方が異なる同一文の音声波形で、両者の比較からタイミングの相違を見ることができる。2つの波形の音韻継続時間長が同じ場合は、直線ab上にフレーム毎の対応曲線、時間伸縮関数は一致することになるが、両図の破線に見られるように、2つの音声波形には時間軸である対角線上にタイミングの違いが確認できる。図3中のbcdはタイミングの違いをさらにパターンとして提示したもので、bは発話速度のゆらぎ、すなわち時間長の伸び縮みを示し、cは同速度による時間的ずれのみを示し、dは発話速度の減少、すなわち時間長の伸びを示すパターンとして表れる。

このように韻律パラメータによって抽出された韻律特徴は、正規化によりパターン化を行い分類処理される。なお、この波形比較型正規化手法の他、今回は説明を省くがZ値などの重要な正規化手法がある。

4.3 韻律の自動抽出

韻律の自動抽出を受け持つ合成モジュールでは、文献^[6]で提案されたToBIの韻律予測法を利用してF0曲線を生成する。上記のToBI記号表現からもF0曲線を予測することができる。Pierrehumbert & Beckmanは論文^[9]でToBI韻律記号表現からF0曲線を生成する(アルゴリズム)方法を提案している。また、平井らは藤崎モデルを通して基本周波数(F0曲線)を予測する方法の有効性を確かめている。このようにToBIを用いてF0曲線の予測は可能である。このF0予測を用いて自動ラベリングを行う。

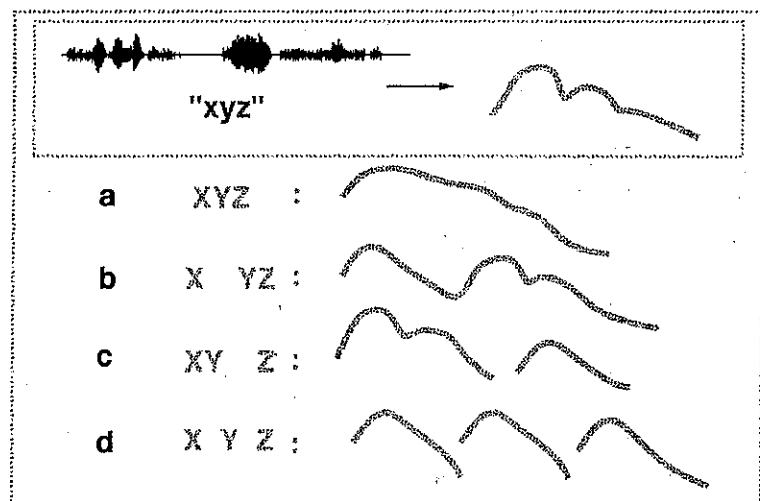


図4 意味解釈のための波形候補選択

図4に意味解釈のために行う波形選択処理、いわゆるセグメンテーションを図式的に示す。出力音声に正しい意味を伝えるための作業、すなわち最適音素の選択は、入力音声から算出したF0曲線と、音声認識部で得られた単語列の情報に基づき行う。まずはじめに単語のすべての可能な読みの組合せを算出する。これには上記の音声合成モジュールから得られたF0予測を通して、それぞれの可能な読みのF0曲線を生成する。次にその中から入力音声に最も近似の曲線が選択される。この作業を自動ラベリングと呼び、この自動抽出された特徴に基づき合成音声生成されるが、さらに強調の違いについても抽出が行われる。図4は音声波形比較による強調抽出を図式的に示す。基準音声と対象音声の2つの波形の変差を捉えることにより強調が認められ、この抽出によって話者の意図や目的意味が出力音声に反映される。

5. 強調と韻律・「意味」の違い

これまで韻律情報処理についてToBI記述を含めて理論的観点から述

べてきた。本節ではさらに具体のレベルに移る。すなわちToBIの韻律記号としての有効性の検証すると共に、韻律と強調の関係について調べる。PI理論とToBIの関連に戻れば、前述の通りL1アクセントによる相違、L2文法的ブラケットング(枝分かれ)と統語論的チャンキングについてはよく符合する。しかし、L3のフォーカル・プロミネンス(または意味的強調)では、ToBIで充分機能しうるかどうか、検討の必要がある。したがって以下、プロミネンス専用データベースを用いて、実験1および実験2により、韻律と強調の関係を明らかにする。

5.1 実験方法と音声データ

使用した音声データは、1名の男性話者が、会議の登録文を複数場面で発声したもので、25文例、各々強調の異なる言い方(2~4種)計89文である。

対話文は、例えば、

- 1 「地下鉄で北大路駅まで行ってください。」
発話1、特に強調を含まない、以下defaultと呼ぶ。
- 2 「地下鉄で北大路駅まで行ってください。」
発話2、強調を含む。
- 3 「地下鉄で北大路駅まで行ってください。」
発話3、強調を含む。

このように同一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。

談話場面が異なれば、当然強調のある箇所も異なる。つまり、発話1の場面は単に一般経路情報を言っている場合、発話2の場面は、バスや電車ではなく、地下鉄で行ってくださいという場合、発話3の場面は、南大路駅などではなく北大路駅までという場合がそれぞれ想定できる。したがってこれらの場面や状況に応じてフォーカス(目的情報)を「強調」の違いによって示すと考えられる。

実験1、実験2の手順については、対比的に図示したい。

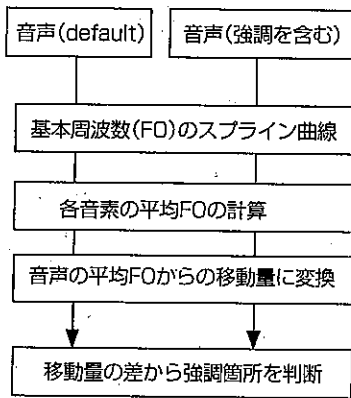


図5 実験1 F0によるフォーカス抽出手法

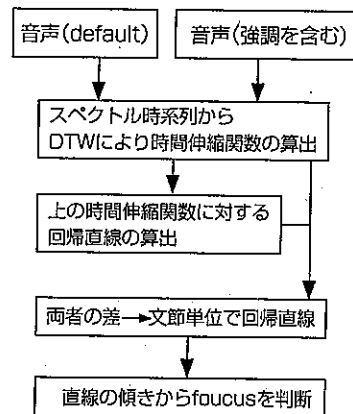


図6 実験2 タイミングによるフォーカス抽出手法

5.2 実験1 (F0とフォーカル・プロミネンス)

実験1では、はじめにF0とフォーカル・プロミネンスの関係を明らかにするために、強調によるピッチの違いを調べる。次にシンボル表記をとるToBI記述法の有効性を調べるために、ToBI予測による基本周波数パターンを用いて同様の実験を行う。実験の方法は、上記の強調の異なる音声データとのF0比較を行うもので、図5の手順に従った。まずはじめにアライメントを行い、F0を滑らかにつなぐ線、いわゆるスプライン曲線を得る。次に各音素ごとにF0の平均値を算出して、さらに平均F0から移動量の差分を求める。なお、フォーカスのある発話単位の判断基準は、1) F0の代表値の差分が最大である音素を含む、2) 発話単位に含まれる全体音素の、F0の差分の平均値が最大、3) その発話単位に含まれる全音素のF0の正の差分の平均が最大、それらの内、今回は1)を用いた。

実験の結果、ピッチ情報により強調を精度よく抽出することができた。図7に「今回は割引をおこなっておりません」を例とした強調を示す。この図では音素を横軸に、差分を縦軸に、特に強調のない音声

を基準発話(default)として縦軸の0にした。実線(focus 1)は「今回は」に上記で得られた差分の最大ピークが、破線(focus 2)は「おこなって」に最大ピーク、つまりフォーカスがある。

このように複数の発話からそれぞれ異なる強調箇所が得られ、基本周波数、つまりピッチのみの違いを用いて強調を示すことができた。したがって、フォーカルプロミネンスとF0には相関関係があることが確認できた。

さらに同様の実験を、CHATR音声合成モジュールのToBI韻律予測を含めて行った^[8]。上記のように実験では同じ話者による強調を含まない発話を基準音声として用意できるが、実際の音声翻訳下では当然ではあるが、比較に用いるための基準音声を前もって用意することはできない。したがって、この場合の基準音声は音声合成モジュールによって生成したものを用いた。つまりこのような方法をとって、ToBI記述を利用した韻律予測を使用することによってToBIの有効性を帰納法的に調べた。

その結果、ToBI予測の場合、誤差が増えるが先の実験とほぼ同程度の7割以上の精度で強調箇所を抽出できた。図8に「今回は割引をおこなっておりません」を例とした抽出の結果を示す。

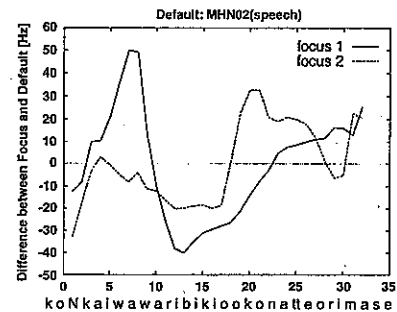


図7 3文の比較による強調箇所の表示
<0破線は強調なし、フォーカス1は「今回は」、フォーカス2は「おこなって」に強調あり。>

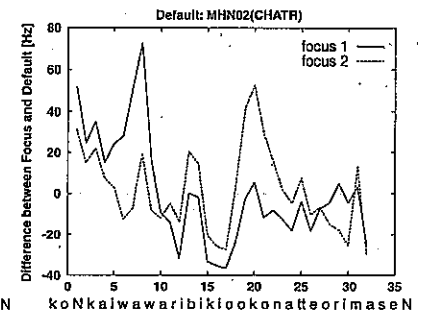


図8 CHATRが予測したF0曲線による強調箇所の抽出
<視覚的には実音声との異なりが見られるが抽出そのものには大差がないことを示す。>

以上、2つの実験結果の比較から、自然発話では81.3%、ToBI予測では75.0%の正確率が得られた。両者には有意の差がなく、強調の抽出にはToBIを用いて予測した情報が有効であることが確かめられた。

5.3 実験2 (タイミングとフォーカル・プロミネンス)

F0とフォーカル・プロミネンスの関係、およびToBIの有効性を確認する実験1に続いて、実験2ではタイミングとフォーカル・プロミネンスの関係について調べる。これまで日本語では、タイミングと強調の関係についての研究は多くは見られない。おそらく日本語はモーラ言語であるという考え方が浸透しており、ストレス言語に見られるような時間伸長は一般に起こらないと考える傾向があるのであろう。しかし韻律はそれぞれの韻律特徴が個別に機能するのではないと考えられることから、以下のように比較実験を行い、両者の関係について明らかにする。

実験の方法は、タイミングを表す音韻継続時間長を用いて、強調の抽出を行い検証する。実験2の手順は上掲、図6に従う。なお、既に説明を加えたDTW音声波形信号比較型正規化手法をこの実験で用いた。

タイミング、すなわち発話速度の違いを明らかにするために、同一文の波形1(強調のない音声)と波形2(強調のある音声)の両波形をDTWにより比較する。音韻継続時間長の差は、図9abcdのほぼ対角線上にのる時間伸縮関数をつなぐ曲線に示される。算出結果は図9ab自然音声との比較、図9cd合成音声との比較によるもので、共にアクセント句ごとの回帰曲線がタイミングの変化を表している。

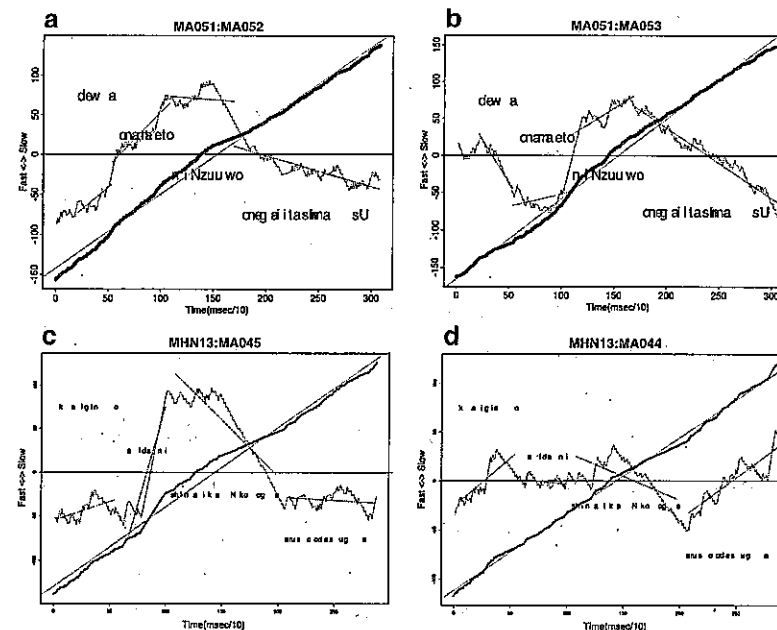


図9 DTWによるフォーカス抽出

ab 「ではお名前と人数をお願いいたします。」
 cd 「会議の間に市内観光があるようですが。」

(ab: 自然音声との比較、aは「お名前」にフォーカスを置き、bは「人数」にフォーカスを置く、タイミングの差を示す時間伸縮関数から得られた結果。cd: 合成音声との比較、cは「間に」にフォーカスを置き、dは「会議の」にフォーカスを置く。)

実験の結果、時間長の伸びの見られる箇所は強調箇所と一致し、60%近い精度で強調箇所を抽出することができた。またランダムチャンスの場合では、20%から30%の間が期待できた。したがって強調は、タイミングすなわち音韻継続時間長と相関関係にあることが分かった。また、他の言語(例えば、英語などのストレス言語)では強調を示すために、音韻継続時間長が長くなることは既に知られているが、日本語においても同様の傾向が見られることを示すことができた。

6. 音声文法の構築

人間の意味理解の解明ではなくても、人間もおそらく上記の比較実験のように韻律をある基準と比べて解釈しているのではなからうか。認識されるのはある基準に対する差であって、その差が解釈にとって重要な意味情報を持つと考えられないだろうか。言い換えればパタン間の差異によって意味解釈が構造化されると思われる。

6.1 韻律はゲシュタルトなものである

上記の実験で並行的にToBIの有効性と、強調と韻律の関係を確かめた。ここで再びToBI記述の観点に立ち戻ることにより、音声文法の基本単位の枠組みについて考えてみる。ToBIは前述の通りPI理論のL3の段階までは示すことができる。しかしL3いわゆる強調については直接記述することはできない。ToBIラベリングそのものは強調を示さない。しかし実験でも用いたように、ToBIのシンボル記号を利用した算出結果を用いて、間接的に記述する方法を考案した。つまりToBIが示すアクセント句ごとのF0の最高値の差異情報に基づいてF0予測アルゴリズムによってL3の強調を示したのである。この間接的記述の方法は、ToBIの韻律付与法がカテゴリの違いを記述するものであって、量的記述は行わないという原則からくる制約が関与している。

ToBIにおける情報の捉え方は二次元記述である。これに対して実際の韻律は、個々の韻律情報の組み合わせが重要な意味を持ち、韻律そのものはゲシュタルトとして、すなわち統合的なものとして機能する。完全とは言えないまでもToBIによってPI理論のL3までの記述は可能であるが、さらに高次の意味解釈を考えた場合、今後加えられるべき情報や記号について検討が必要となる。

6.2 今後の課題

上記の実験に関する直接的課題として、F0とタイミングの関係は依存的に意味を持つものか、あるいはそれぞれ独立に意味を持つものか、

また時間長の変化がF0の影響からくるものであるか、さらに検討を要する。

韻律記述に関しては、ToBI記述法は比較的确实な手法であり、多くの言語に適用され始めている。しかし、意味を持つ韻律パタンの差を如何に記述するかという問題については、F0とブレイクインデックスだけを基本情報とするのではなく、声の強さ、声の質のカテゴリ化も考えられる。

7. むすび

本稿では、音声文法構築への第一歩と考え、音声情報処理の観点から韻律と「意味」の関係を踏まえ韻律記述の方法について検討した。すなわち、ToBI韻律付与法の有効性を検証するとともに、その記述能力を韻律と「意味」との関係から調べた。

また強調あるいはフォーカル・プロミネンスについて、同じ単語列であっても複数の韻律情報を用いることによって話者意図や目的意味を示す強調を抽出できることを明らかにした。

第一に声の抑揚に対しては、default音声との比較により強調を示すことができると分かった。すなわち、ここでは従来のToBI手法だけで付与できないテキストからの基準ToBIパラメータを通して、ピッチ曲線を予測したパタンと入力音声から抽出したピッチ曲線のパタンとの比較を行い、プロミネンスの位置を判断するという新たな手法を用いた。

第二に音韻継続時間長に対しては、声の高さいわゆるピッチと共に、タイミングが同様の影響を持つことが分かった。これまでの研究では、ストレス言語でない日本語に対して音韻時間長についての考察はあまり加えられなかった。しかし実験の結果にも見られるように音韻継続時間長に伸長が認められた。したがって、モーラタイミング言語と考えられている日本語においても、音韻時間長伸長にみるタイミングは韻律情報として重要であると思われる。

さらに高次の意味解釈を考えた場合、上述のL5の感情などをどのような記号で付与すればよいのか、検討の必要が認められる。しかし、現在の技術と合わせることによって、L3までの意味情報抽出にはToBIは充分その機能を果たし、言語的意味の違いは伝えることができる。ただし、韻律によって示されるものは、現在自動的に測定あるいは抽出可能でないものを含めると、さらに多くの情報が存在する。それらに対して、何が音声文法の主要な決め手となりうるかを決定・判断するかが今後の大きな課題である。韻律の基本単位の決定はもとより、韻律の大規模データベースの構築はその文法的枠組みの決定に多くの指針を提示できるであろう。

謝辞

I would like to express my deep gratitude to Kumiko Hayakawa-Campbell for her help with the writing of this paper, and to Satoshi Kitagawa for his help with the experiments.

参考文献

- [1] キャンベル, ニック (1997) 「プラグマティック・イントネーション: 韻律情報の機能的役割」『文法と音声』音声文法研究会編(くろしお出版).
- [2] M. E. Beckman & G. M. Ayers (1993) "The ToBI Handbook", Tech. Rept., Ohio-State University, U.S.A.
M. E. Beckman & G. M. Ayers (1994) "Guidelines to ToBI Labelling". Version 2.0., February.
- [3] G. Bruce (1977) "Swedish Word Accents in Sentence Perspective", Lund, CWK Gleerup Laromedel.
- [4] W. N. Campbell (1996) "Autolabelling Japanese ToBI" Proc. ICSLP-96 (Philadelphia) pp. 2399-2402.
- [5] W. N. Campbell (1997) "The ToBI system and its application to Japanese" pp. 223-229, Journal of the Acoustical Society of Japan 53, 3, (in Japanese).
- [6] M. Grice, M. Reyelt, R. Benzmueller, J. Mayer & A. Batliner (1996) "Consistency in transcription and labelling of German intonation with G-ToBI", Proc. ICSLP-96, pp. 1716-1719.

- [7] J. J. Venditti (1995) "Japanese ToBI Labelling Guidelines", Tech. Rept., Ohio-State University, U.S.A.
- [8] W. N. Campbell & A. W. Black (1996) CHATR: 自然音声波形接続型任意音声合成システム, 電子情報通信学会技術研究報告 SP96-7 (May '96) pp. 45-52.
- [9] J. B. Pierrehumbert & M. Beckman (1988) "Japanese Tone Structure", Cambridge, MA: MIT Press.