

日本語単語アクセントの基本周波数パターンとその生成機構のモデル

藤 崎 博 也・須 藤 寛

(東京大学工学部)

(昭和46年6月5日 受理)

A Model for the Generation of Fundamental Frequency Contours of Japanese Word Accent

Hiroya FUJISAKI and Hiroshi SUDO

(Faculty of Engineering, University of Tokyo)

(Received June 5, 1971)

Prosodic features in speech can be interpreted as responses of the underlying mechanisms to a set of linguistic commands. This paper presents a quantitative model for the mechanisms of generating fundamental frequency contours of word accent of standard Japanese.

All the types of word accent of standard Japanese are characterized by the existence of a transition in the subjective pitch, either upward or downward, at the end of the initial mora, and by the fact that no more than one downward transition is allowed within a word. Table 1 lists the patterns of subjective pitch of all the possible accent types of words that consist of up to 5 morae.

These binary patterns, however, never manifest as such in the fundamental frequency contours. Analysis of utterances of a number of speakers (Fig. 1) indicates that the logarithmic fundamental frequency contours of the same word accent, when normalized both in time and in frequency, are essentially similar (Fig. 2 and Fig. 3). These observations lead to the model of Fig. 4 based on the following assumptions:

- (1) Each type of word accent can be characterized by a unique logarithmic contour.
- (2) Commands for voicing and accent take the form of binary input to the system.
- (3) Separate mechanisms exist for voicing and accent, which can be approximated by linear systems that convert the binary commands into the respective control signals (Fig. 5).
- (4) These control signals are combined and applied to the mechanism of glottal oscillation, whose fundamental frequency is an exponential function of the control signal.
- (5) The glottal mechanism shows hysteresis specified by the onset and cessation of the oscillation (Fig. 6).

In order to investigate the validity of the model, fundamental frequency contours of various utterances of isolated words were extracted by a computer program (Fig. 7) and were analyzed by the method of Analysis-by-Synthesis (Fig. 8). A few examples of the comparison of the extracted fundamental frequency contour and its closest approximation obtained by the A-b-S procedure are shown in Fig. 9.

1. はじめに

最近の音声合成手法の進歩は、音声生成の音響的ないし調音的な過程の観測・分析とそのシミュレーションに基礎をおいたものであり、いわゆる規則による合成¹⁾、すなわち離散的な記号系列から、明瞭性、自然性ともにすぐれた連続的な音声信号を合成する可能性を開いている。とりわけ、音韻性に寄与の大きい声道の音響的特性に関しては周知のように回路理論的な手法を用いて、これを比較的少数のパラメータによってかなり正確に記述しうることが明らかにされている。一方、いわゆる韻律性は、主として声帯音源の基本周波数、強度、および音素の持続時間等のパラメータによって表現される。しかしながら上述の規則による合成の実現に必要なのは、入力記号系列と、これらのパラメータの時間的変化パターンとの間の本質的な関係を簡潔で的確な法則によって記述することであり、英語および日本語の連続音声の規則による合成に関する在来の研究²⁻⁶⁾は、かなりの明瞭度を達成し得ているとはいえ、なお完成の域に達したとは言い難い現状であって、特に、韻律性の合成にはなお多くの改良の余地が残されているようにおもわれる。

韻律的特徴の主要な音響的表現としての声帯音源の基本周波数は、単語アクセント、強調、イントネーション、構文などの言語学的情報を担うと同時にこれらの情報が個人の声帯振動機構を介して実現される過程において、話者の個人性、および発話時の情緒などの非言語情報をも賦与されている。しかしながら、規則による合成の見地からは、言語学的情報を基本周波数の時間的変化に変換する過程を定量的に表現することが最も重要である。この問題に関する従来の研究を概観すると、S. Öhman^{7,8)}は2種類の言語学的指令を入力とする発声機構のモデル化を試み、これによりスウェーデン語の多くの方言のイントネーションの統一的な説明を行なっている。また、比企ら⁹⁾は各種の要因の複合による基本周波数の変動パターンの表現を試みた。I. Mattingly³⁾は3種の言語学的要因から基本周波数パターンを折線により合成する規則を発表している。同様の試みは斎藤ら¹⁰⁾、松井ら⁴⁾によっても行なわれた。一方、中山ら¹¹⁾はモーラ内平均基本周波数を指数関数的な自然降下と、分節単位内での上昇・降下との和として表現し、それらを線分で連結することを試みてお

り、また加藤ら¹²⁾は瞬時基本周波数を滑らかな曲線で近似する方法を示している。また L. Rabiner⁵⁾も文章の基本的な台形パターンに強調の位置を示す部分的な起伏を重畳することによって基本周波数パターンの合成を試みている。

このように在来の研究は規則による合成の進歩に貢献するところが少なくなかったとはいえ、少数の例外を除いては比較的単純な規則の導出を主眼としたため、言語学的情報から基本周波数の時間的変化への変換の過程および種々の言語学的情報の間の関連にまで深く溯ったものが乏しく、現実には観測される基本周波数の時間的変化パターンを十分な精度で近似し得るに到っていない。上記の S. Öhman の研究はこの点では極めて異色なものであるが、その採用したモデルにはなお検討すべき点が少なくない。本論文ではこのような見地から、特に標準日本語の単語アクセントを表わす基本周波数パターンの生成を目的として、まず現実には観測される基本周波数パターンを分析し、その結果に基づいて音声の基本周波数制御機構の機能的なモデルを提案し、さらにこのモデルに立脚し、観測される種々の基本周波数パターンが高い精度で近似できることを示したものである。なお、基本周波数の時間的変化は上述の言語学的要因のみでなく、音素、特に子音の存在によって影響を受けることは周知の事実であるが、その影響は前記の諸要因によるものと比べれば比較的局所的なものであるので、本研究ではこれらを度外視し、基本周波数の時間的変化の巨視的な傾向を対象とすることとした。

2. 単語のアクセント型とその基本周波数パターン

標準日本語（東京方言）の単語アクセントの特徴は第1モーラから第2モーラにかけて必ず著明なピッチの上昇または下降があり、かつ一単語内でのピッチの著明な下降は一ヶ所のみに限られる。したがって、 n 個のモーラから成る単語には $(n+1)$ 種のアクセント型が存在する。5モーラ以下の単語についてこれを図式的に分類すると、Table 1の通りである¹³⁾。表中・印は n モーラ単語の各モーラのピッチを、○印は後続の助詞のピッチを表わす。以下本論文では、第 i モーラと第 $(i+1)$ モーラの間でピッチが著明に下降するものを (n, i) 型の単語アクセントと名づけ、また、著明な下降を有しないものを $(n, 0)$ 型と名

Table 1 Patterns of subjective pitch of all the types of word accent in standard Japanese for words that contain up to five morae.

		Number of morae in a word				
		1	2	3	4	5
Types of word accent	0	i	o-i	a-o-i	o-i-o-i	ni-ho-n-ga-mi
	1	o	a-o	a-wa-re	o-i-o-i	a-me-a-ra-re
	2		i-e	a-o-i	i-e-i-e	a-o-i-u-o
	3			u-re-i	a-o-a-o	a-ma-ga-e-ru
	4				o-to-o-to	mj-a-ya-ma-ru
	5					o-sho-o-ga-tsu

づける。

以上のように分類されたアクセント型は主観的なピッチの高低を抽象的に表現したものであるが、現実の基本周波数の時間的変化のパターンは、一般に時間の連続関数であって、上記のように単純な波形を示さない。したがって、音響的現象としての基本周波数パターンと主観的なピッチパターンおよびその言語学的な情報との関係を客観的に把握するには、まず実際に発声された単語の基本周波数を測定することが必要である。本研究ではこの目的のために、なるべく連続的で明瞭な基本周波数パターンが得られるよう、モーラ数5以下のすべてのアクセント型の単語のうち、主として母音・半母音等の有声音から成るもの (Table 1 に示す) を選び、これらの単語を、単独に、また“それは——というものです”という文章に挿入して発声したものを測定の対象とした。発声者は男女各3名であり、発声速度は特に規定せず、自然な発声にまかせた。この段階での基本周波数パターンの測定にはソナグラフを用いたが、発声者の個人差を除去して個々のアクセント型に固有の基本周波数パターンを見出すには、まず第一に時間領域での正規化によって発声速度の個人差を除去することが必要である。そこですべての発声サンプルから、1モーラの平均持続時間 (この場合 0.12 sec) を求め、実測された基本周波数パターンの持続時間がすべてこの値の整数倍となるように時間軸を正規化した。Fig. 1 は前記6

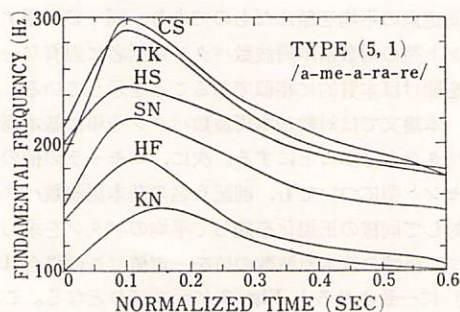


Fig. 1 Measured fundamental frequency contours of utterances of the sample word /a-me-a-ra-re/ by six speakers.

名が (5, 1) 型の単語“アメラレ”を単独に発声した場合の基本周波数パターンの時間軸をこのように正規化して示したものであるが、共通のパターンを見出すにはさらに周波数領域での正規化が必要なることは明らかであろう。その目的のために、得られた基本周波数の実測値を対数値に変換し、さらに、各パターンの平均基本周波数が 120 Hz になるように移動してプロットすると Fig. 2 に示すようになる。図中の曲線は、

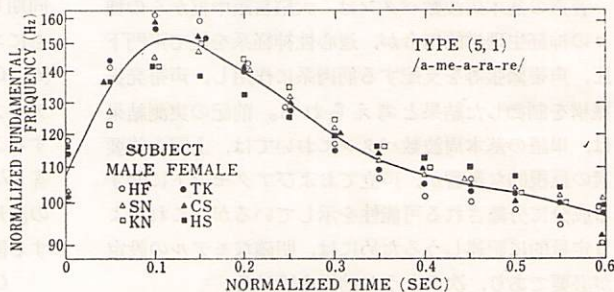


Fig. 2 Normalized logarithmic fundamental frequency contours of utterances of the sample word /a-me-a-ra-re/ by six speakers.

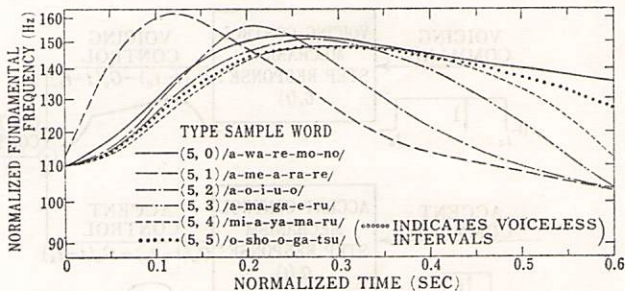


Fig. 3 Average logarithmic fundamental frequency contours for all the accent types of 5-mora words.

各測定点の平均を結んだものであり、同一の単語アクセント型の対数基本周波数パターンが話者に固有な一定値を除けば本質的に相似であることを示している。以下、本論文では対数基本周波数パターンを単に基本周波数パターンと呼ぶことにする。次に、5モーラの他のアクセント型についても、前記6名の基本周波数パターンに対して同様の正規化を施して平均のパターンを求め、かつ、語頭の基本周波数の値を一定値（この場合 110 Hz）に一致させると Fig. 3 に示す通りとなる。これらの曲線は、すべてのアクセント型の基本周波数パターンが、アクセント型によらず語頭でやや急激に上昇したのち語尾に向かって緩やかに下降する成分と、アクセントの位置に対応して滑らかに上昇したのち下降する成分との和として近似的に表現し得る可能性を示しており、前者はすべての単語の発声に必要な声立ての結果であり、後者がアクセントの位置を示すための発声中の基本周波数の調節の結果であると考えられることができる。

3. 単語のピッチ制御機構のモデル¹⁴⁾

音声の基本周波数パターンは、大脳言語中枢からの種々の神経生理学的指令が、遠心性神経系を経て声門下圧、声帯緊張等を支配する筋肉系に作用し、声帯発振機構を制御した結果と考えられる。前記の実測結果は、単語の基本周波数パターンにおいては、上記の諸要素の巨視的な影響が、声立ておよびアクセントに関する成分に分離される可能性を示しているが、これをより定量的に記述しうするためには、明確なモデルの設定が必要であり、次のような前提を設ける。

(1) 同一アクセント型の単語の基本周波数パターンは、発声者の平均基本周波数の差、および発声速度の

差を除いては本質的に同一の時間的変化を示す。

(2) 声立ておよびアクセント等の言語学的特徴を実現するためのモデルへの入力指令はすべて2値的な信号とする。

(3) 声立ておよびアクセントに関しては、それぞれの別個の制御機構があり、前記指令に応じてピッチ制御信号を生成するが、それらの特性は近似的に線型の表現が可能である。

(4) 上記の声立ておよびアクセントの制御信号は相加されて声帯振動機構に加えられ、その出力として基本周波数が定まるが、その間の変換は指数関数で近似できる非線型を有する。

(5) 声帯の振動機構は、一般にヒステリシス特性を有し、振動開始時の基本周波数と停止時の基本周波数とは異なる。

上記の前提に立脚した単語のピッチ制御機構のモデルを Fig. 4 に示す。声立ておよびアクセントに関する指令は単位ステップ関数とし、前者は発声開始の直前から停止の直前まで連続的に加えられ、後者は単語のアクセント型に対応して特定のモーラの発音とはほぼ同期して加えられる。それぞれの制御機構の特性はともにステップ応答で記述するものとし、それぞれ $G_v(t)$ 、 $G_a(t)$ と表現する。前節の考察から、声立てのステップ応答 $G_v(t)$ はいったん上昇した後、次第に下降するような時間波形であり、アクセントのステップ応答 $G_a(t)$ は $t=0$ で連続な微係数を有する単調非減少の波形と考えるのが妥当であろう。これらの特徴を有する関数のうち、最も簡単なものはそれぞれ、

$$G_v(t) = A_v a t e^{-\alpha t}, \quad G_a(t) = A_a [1 - (1 + \beta t) e^{-\beta t}]$$

である。ただしここで A_v 、 A_a はそれぞれの応答の大きさを表わす振幅パラメータ、 α 、 β は応答速度を

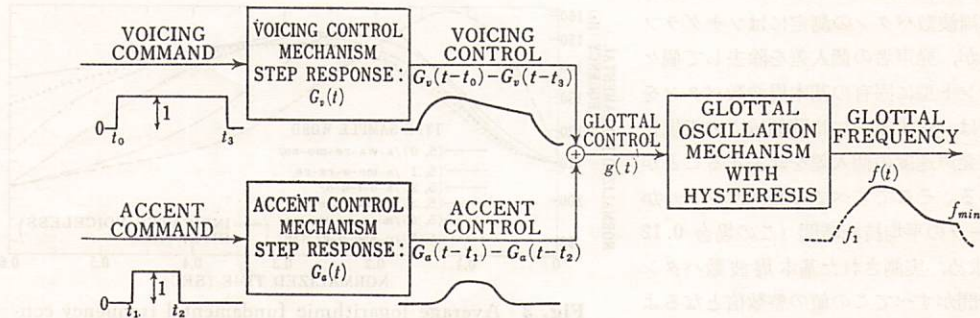


Fig. 4 A model of control mechanisms of glottal frequency.

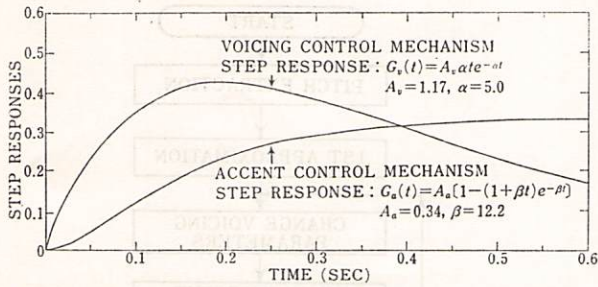


Fig. 5 Step responses of control mechanisms for voicing and accent derived from measured logarithmic fundamental frequency contours.

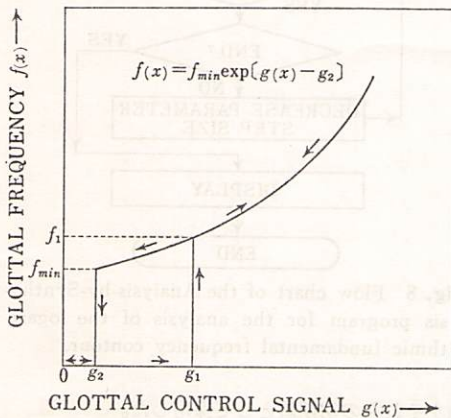


Fig. 6 Nonlinear characteristic of the glottal oscillation mechanism.

表わす時間パラメータである。Fig. 5 はこれら二つの関数の時間波形を示したものである。

一方、これらの制御信号を現実の基本周波数に変換する声帯振動機構の特性は Fig. 6 に示すごとくであり、発振開始および停止に相当する入力信号の値をそれぞれ g_1 および g_2 とすると、出力としての基本周波数 $f(x)$ は、入力信号 x が増加する場合には

$$f(x) = 0 \quad (g(x) < g_1 \text{ のとき})$$

$$f(x) = f_{\min} \exp(g(x) - g_2) \quad (g(x) > g_1 \text{ のとき})$$

となり、逆に入力信号 x が減少する場合には

$$f(x) = f_{\min} \exp(g(x) - g_2) \quad (g(x) > g_2 \text{ のとき})$$

$$f(x) = 0 \quad (g(x) < g_2 \text{ のとき})$$

で与えられる。ただし f_{\min} は発振可能な最低周波数である。したがって上記モデルに基づいた対数基本周波数パターン $f_0(t)$ は次式で近似できる。

$$f_0(t) = f_{\min} \cdot \exp[G_v(t-t_0) - G_a(t-t_3)]$$

$$+ G_a(t-t_1) - G_a(t-t_2) - g_2]$$

ただしここで t_0, t_3 はそれぞれ声立て指令の始点および終点であり、 t_1, t_2 はアクセント指令の始点および終点を表わす。

4. ピッチ抽出と A-b-S 法による分析¹⁵⁾

4.1 ピッチ抽出

上記のモデルの妥当性を検討するに当たっては、まず実際の音声の基本周波数パターンを正確に抽出する必要がある。この問題の解決には種々の手法が発表されているが、ここでは最近有力な手段として提案されているケプストラム法¹⁶⁾をさらに改良して用いた。周知のように、信号 $f(t)$ の時刻 t_0 における電力スペクトルは

$$p(\omega, t_0) = \left| \int_{-\infty}^{\infty} f(t) \cdot h(t-t_0) e^{-j\omega t} dt \right|^2$$

で与えられる。ただし、 $h(t)$ は観測時間窓関数を表わす。ケプストラムは対数電力スペクトルをさらにフーリエ変換したものであり、時間の次元をもつ quefrequency λ の関数として次式のように定義される。

$$c(\lambda, t_0) = \left| \int_{-\infty}^{\infty} h(\tau) \cdot p(\tau, t_0) e^{-j\lambda \tau} \cdot d\tau \right|^2$$

周期的な時間波形の対数電力スペクトルは、周波数に関する周期的成分と、その包絡線成分の和とみなすことができるが、後者の周波数に対する変化は前者と比べて一般に緩やかであるので、ケプストラム上では両者の分離が比較的容易となる。すなわち、考えられる基本周期の範囲に対応した quefrequency の範囲でのケプストラムのピークの位置を検出すれば、それが直ちに基本周期を示すことが多い。しかしながら、基本周期の整数倍の位置に、より大きなピークが生ずる場合もあり、単純な最大値検出を行なうと誤動作を招くので、本研究では quefrequency の平方根に逆比例する荷重を加えた荷重ケプストラム

$$c'(\lambda, t_0) = \lambda^{-1/2} c(\lambda, t_0)$$

を採用し、500 Hz から 67 Hz までの基本周波数に対応して λ の 2 msec から 15 msec までの範囲での最大値の検出を行ない、その位置の quefrequency の値を基本周期とした。この方法で抽出される基本周期が意味をもつのは、当然有声区間に対してのみであるから、別の基準にもとづいてあらかじめ有聲と判定された区間についてピッチ抽出を行なった。その結果の1例を

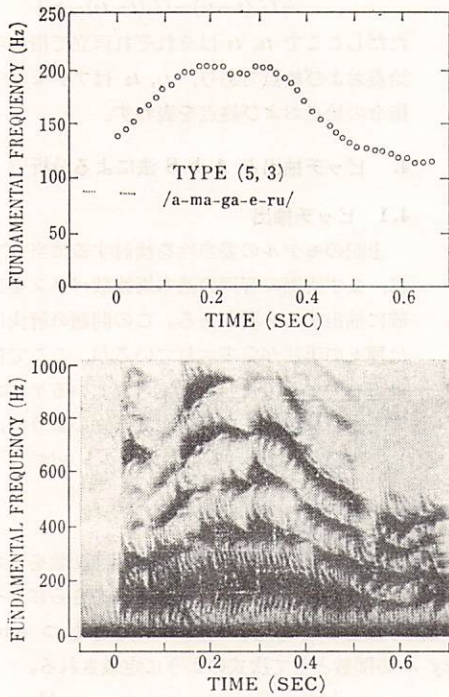


Fig. 7 The fundamental frequency contour of a word /a-ma-ga-e-ru/ obtained in every 16 msec by the pitch-extraction program (above), and the corresponding narrow-band sound spectrogram (below).

狭帯域ソナグラムと対照して Fig. 7 に示す。

4.2 A-b-S 法による基本周波数パターンの分析

上記の手法により抽出した基本周波数パターンからピッチ制御機構モデルのパラメータを決定するには A-b-S 法¹⁷⁾によった。A-b-S の手法に関しては周知であるのでここでは概略のフローチャートを Fig. 8 に示すにとどめるが、逐次近似法であるので第 1 次近似としてモデルに与えるべきパラメータの初期値の適否が分析の精度や能率を左右する。モデルのパラメータの中、 A_v , A_a , α , β は主として個人差により変動するが、分析の対象とした発声の範囲ではそれらの初期値としてはそれぞれ 1.2, 0.35, 5.0/sec および 12/sec が妥当であった。また声立ておよびアクセント指令の変化時点 $t_0 \sim t_3$ は、単語の長さおよびそのアクセント型に依存するので、それぞれに固有の初期値を与えることが分析の能率上望ましい。(n, i) 型の単語について発声開始時点を時間の原点にとった場合、以下の

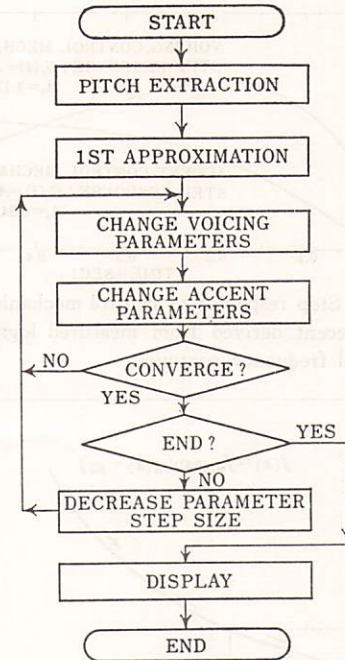


Fig. 8 Flow chart of the Analysis-by-Synthesis program for the analysis of the logarithmic fundamental frequency contour.

数値をそれらの初期値として採用した。

$$t_0 = -0.05 \quad \text{sec}, \quad t_3 = n \cdot T_m \quad \text{sec},$$

$$t_1 = \begin{cases} -0.05 \quad \text{sec} & (i=1 \text{ の時}), \\ (T_m - 0.05) \text{ sec} & (i \neq 1 \text{ の時}), \end{cases}$$

$$t_2 = \begin{cases} t_3 \quad \text{sec} & (i=0 \text{ の時}), \\ (i T_m - 0.05) \text{ sec} & (i \neq 0 \text{ の時}), \end{cases}$$

ただしここで T_m は 1 モーラの平均持続 (0.12 sec) 時間である。分析の手順はフローチャートに示すように、まず声立ての制御パラメータ (t_0 , t_3 , A_v , α), 次にアクセントの制御パラメータ (t_1 , t_2 , A_a , β) の順に変化させて逐次近似を行なう。モデルと実測値との整合の評価には有声区間における平均自乗誤差を用いた。この方法で最大の精度を得るためには、すべてのパラメータの値が完全に収束するまで計算を行なうことが望ましいが、分析の速度を上げるため、ここでは近似を第 5 次までで打切った。Fig. 9 (a)~(d) は 5 モーラの単語の基本周波数パターンを分析した若干の例を示すもので、 \circ 印はケプストラム法による基本周波数の実測値、曲線は A-b-S 法による第 5 次近似で

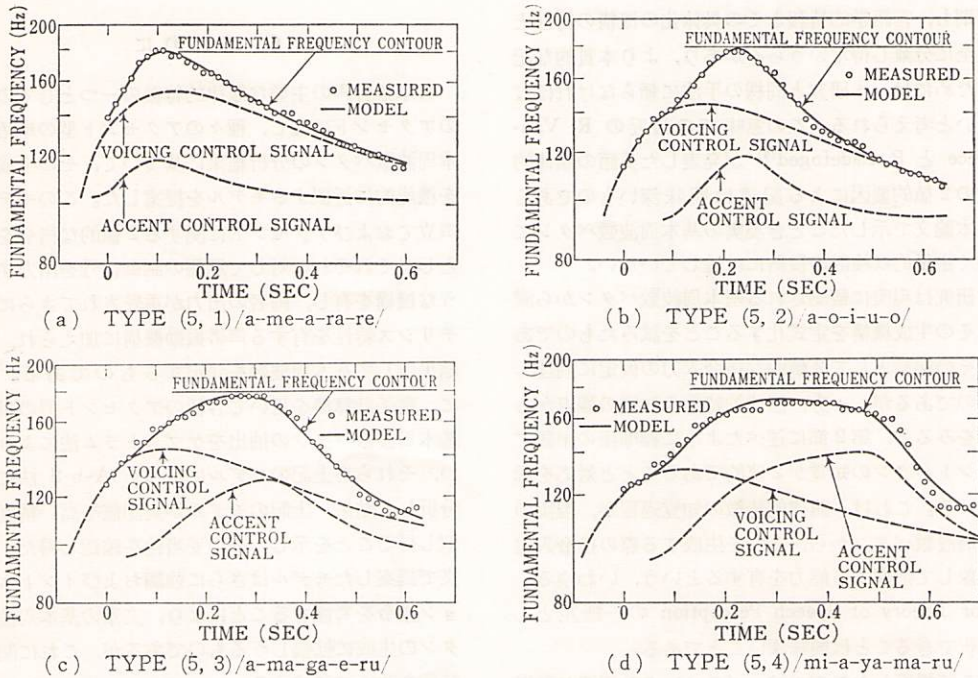


Fig. 9 Examples of analysis-by-synthesis of word fundamental frequency contours.

あり、実測値とモデルとの相対誤差は $\pm 4\%$ 以下にとどまっている。

5. 考察および検討

すでに本論文の冒頭で述べたように、韻律的特徴の生成あるいは近似に関する従来の研究の中で、S. Öhman の研究は最も基本的な観点に立脚したものであり、単一のモデルを用いてスウェーデン語の多数の方言の韻律的特徴を統一的に説明しうることを示した点で、生成モデルによる表現が現象の本質的把握のために極めて強力な手段であることを立証している。しかしながら音声の基本周波数パタンの生成機構のモデル化の真の意義は、現実に観測されるパタンの呈する特徴の中、呼吸・発声機構の有する生理的・物理的な拘束に由来するものをそれらの特性として定量的に記述することによって、言語学的な情報と、基本周波数の生成機構とを完全に分離し、音声現象とその言語学的内容との間の最も直接的かつ普遍的な対応関係を見出すことにあり、この意味では、Öhman のモデルは必ずしも十分な成功を収めているとは言いがたい。すなわち彼のモデルでは、すべての発話に共通に観測される

基本周波数の話頭での上昇とそれに続く自然降下とを自動的に表現しうる記述を採用していないために、言語学的情報と生理・物理的機構の影響とを完全に分離し得ず、その結果、入力指令として任意のレベルの多値の階段関数を必要としている。第3節で導入した前提 (2), (5) および声立て制御機構のステップ応答はこの見地から着想されたものであり、それによって我々のモデルでは、言語学的な入力情報と基本周波数パタンの生成機構の特性の両者をほとんど完全に分離し、それぞれに関する本質的な記述を得たものと考えられる。特に韻律的特徴の生成のための2値的指令の意味する所は大である。なお、われわれのモデルは、声立て制御機構が、一定の指令に対しても一定の出力を維持し得ないような筋肉・空気力学系（たとえば呼吸圧制御系）によって実現されるのに対し、アクセント制御機構は、一定の指令に対して一定の状態を維持することのできる筋肉系によって実現されていることを示唆するものである。

この関連において英語の主要な韻律的特徴としてのストレスパターンに関する N. Chomsky と M. Halle の最近の試論¹⁸⁾に言及すれば、彼らの提案している生

成規則も、言語学的情報とその具体化の機構の特性とを完全に分離し得ないうらみがあり、より本質的な記述のためには、本研究と同様の手法に頼らなければならないと考えられる。この意味での最近の R. Vanderslice と P. Ladefoged¹⁹⁾ が発表した英語の韻律的特徴の2値的要因による記述は興味深いものであるが、本論文で示したごとき現実の基本周波数パターンに基づく定量的な検討の段階には達していない。

本研究は現実を観測される基本周波数パターンから溯ってその生成機構を定式化することを試みたものであり、その帰結として2値的な指令入力の仮定に到達したものであるが、一方、韻律的特徴の知覚の観点からこれをみると、第2節に述べたように標準語の単語アクセントパターンの知覚が2値的であることと対応を示している。これは、韻律的特徴の知覚過程が、現実の基本周波数パターンから、それを生成する際の指令入力を抽象して把握する能力を有するという、いわゆる、Motor Theory of Speech Perception の一証左として解釈できることは興味深いことである。

ここで提案したモデルは、パラメータの適切な選択によって、現実の基本周波数パターンを高い精度で近似しうるものであり、それによって前述の諸仮定の妥当性を立証しているが、このようにして求められるパラメータの値は、話者の個人差により、また同一話者においても発声ごとに多少の変動を免れないのは当然である。一方、規則による合成の見地からは、それらのパラメータの値および相互の関係を、合成音声の自然性を損うことなしにどの程度まで簡略化し得るかが問題であり、この観点から、ターミナルアナログ方式の電子計算機シミュレーションにより、合成および試聴による実験的検討をすでに行なったが²⁰⁾、本論文の範囲を出るので割愛する。また本論文では、対象を単語アクセントの基本周波数パターンに限定したが、ここで導いたピッチ制御機構のモデルは、強調およびイントネーションに関しても2値的な指令を仮定することによって、文章の基本周波数パターンの生成にまで拡張できることを示した²¹⁾。さらに従来の単語および呼気段落に代る単位としてアクセント単位、および発声区間を導入し、それらの単位の形成と構文構造との関係を明らかにしたが²²⁻²⁴⁾、これらについても稿を改めて詳述したい。

6. おわりに

日本語音声の主要な韻律的特徴の一つとしての単語のアクセントに関し、種々のアクセント型の単語の基本周波数パターンの分析結果に基づいて、その生成過程を機能的に近似するモデルを提案した。このモデルは声立ておよびアクセントに関する2値的な指令を入力とし、それぞれに対して別個の制御信号を出力するような機構を有し、両者の出力が重畳されてさらにヒステリシス特性を有する声帯振動機構に加えられ、その結果として基本周波数を決定するものである。つぎに、電子計算機を用いて各種のアクセント型の単語の基本周波数パターンの抽出をケプストラム法により行ない、それらを上記のモデルに基づくA-b-S法により分析した結果、上記のモデルが実測値を高い精度で近似し得ることを示し、その妥当性を確認し得た。本論文で提案したモデルはさらに強調およびイントネーション指令を考慮することにより、文章の基本周波数パターンの生成に拡張しうるものであるが、これに関しては稿を改めて詳述する。

文 献

- 1) Cooper, F. S., Liberman, A. M., Lisker, L. and Gaitenby, J. H.: Proceedings of the Speech Communication Seminar, (Stockholm, 1962) F-2.
- 2) Holmes, J. N., Mattingly, I. G. and Shearme, J. N.: Language and Speech, 7, 127-143 (1964).
- 3) Mattingly, I. G.: Language and Speech, 9, 1-13 (1966).
- 4) Matsui, E., Suzuki, T., Umeda, N. and Omura, H.: Reports of 6th ICA, B-5-3 (1968).
- 5) Rabiner, L.: Bell Syst. Tech. J., 37, 17-37 (1968).
- 6) Flanagan, J. L., Coker, C. H., Rabiner, L. R., Schafer, R. W. and Umeda, N.: IEEE Spectrum, 7, 22-45 (1970).
- 7) Öhman, S. and Lindqvist, J.: STL-QPSR-4/1965, 1-6.
- 8) Öhman, S.: STL-QPSR 2-3/1967, 20-54.
- 9) 比企, 山浦, 大泉: 昭和41年度電気四学会連合大会講演論文集, 1884.
- 10) 斎藤, 橋本, 脇田: 昭和42年度音響学会春季研究発表会講演論文集, 2-4-13.
- 11) 中山, 市川, 三浦: 昭和43年度音響学会春季研究発表会講演論文集, 1-2-21.

- 12) Kato, Y., Ochiai, K. and Azami, S.: Reports of 6th ICA, B-5-13 (1968).
- 13) 金田一: 明解日本語アクセント辞典, 三省堂 (1958).
- 14) Fujisaki, H. and Nagashima, S.: Ann. Rep. Eng. Res. Inst., Fac. Eng., University of Tokyo, **28**, 53-60 (1969).
- 15) 藤崎, 須藤, 長島: 昭和44年度音響学会秋季研究発表会講演論文集, 3-2-8.
- 16) Noll, A. M.: J. Acoust. Soc. Am., **41**, 293-309 (1967).
- 17) Bell, C. G., Fujisaki, H., Heinz, J. M., House, A. S. and Stevens, K. N.: J. Acoust. Soc. Am., **33**, 1725-1736 (1961).
- 18) Chomsky, N. and Halle, M.: The sound pattern of English, Harper and Row (1968).
- 19) Vanderslice, R. and Ladefoged, P.: UCLA Working Papers in Phonetics No. 17, 6-21 (1971).
- 20) 藤崎, 須藤: 昭和45年度音響学会春季研究発表会講演論文集, 2-2-15.
- 21) Fujisaki, H. and Sudo, H.: Ann. Rep. Eng. Res. Inst., Fac. Eng., University of Tokyo, **29**, 215-221 (1970).
- 22) 藤崎, 須藤: 昭和45年度電子通信学会全国大会講演論文集, S. 3-5.
- 23) 藤崎, 須藤: 昭和45年度音響学会秋季研究発表会講演論文集, 1-1-14.
- 24) 藤崎, 須藤: 昭和46年度音響学会春季研究発表会講演論文集, 1-1-19.