ヒューマンインタフェース 97−12
音声言語情報処理 40−12
（2002. 2. 1）

マルチモーダルコミュニケーションのための音声合成プラットホーム

山下洋一(*1)，喜多竜二(*2)，峯松信明(*2)，吉村貴克(*3)，徳田恵一(*3)，
田村正統(*4)，益子貴史(*4)，小林隆夫(*4)，広瀬啓吉(*2)

1) 立命館大学，〒 525-8577 滋賀県草津市野路東 1-1-1
2) 東京大学，〒 113-8656 東京都文京区本郷 7-3-1
3) 名古屋工業大学，〒 466-8555 名古屋市昭和区御器所町
4) 東京工業大学，〒 226-8502 横浜市緑区長津田町 4259

あらまし　人間と機械とがマルチモーダルな情報交換を行う対話システムでは，音声合成が重要な要素技術となる．本報告では，現在開発中の音声合成システムの基本構成，外部インタフェース，発話文記述方法，音声合成エンジン，アクセント結合処理について紹介する．
キーワード：音声合成，対話，テキスト解析，韻律生成，HMM，アクセント結合

A Platform of Speech Synthesis for Multimodal Communication

Yoichi Yamashita(*1), Ryuji Kita(*2), Nobuaki Minematsu(*2),
Takayoshi Yoshimura(*3), Keiichi Tokuda(*3), Masatsune Tamura(*4),
Takashi Masuko(*4), Takao Kobayashi(*4), Keikichi Hirose(*2)

1) Ritsumeikan University, 1-1-1 Nojihigashi, Kusatasu, Shiga, 525-8655
2) University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656
3) Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, 466-8555
4) Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, 226-8502

abstract　Speech synthesis is a key technique for multi-modal communication between man and machine. This paper describes a speech synthesis system under development, focusing on the component structure, the interface, the description scheme of utterances, the speech synthesizer, and the accent processing.
keywords: speech synthesis, dialogue, text analysis, prosody generation, HMM, accent processing

1　はじめに

音声対話システムを構成するための音声合成システムには，様々な韻律や声質での音声合成，韻律や声質の外部からの制御，発話を中断する出力制御などの機能が求められる．顔画像出力も伴うマルチモーダルな出力生成では，さらに，他出力モジュールとの同期も必要となる [1]．従来の音声合成の研究は，音声によるテキストの流暢な読み上げを目標として精力的に行われてきており，音声合成を用いた音声対話システムを構築しようとしたときに，使い勝手の良い音声合成システムは見当たらないのが現状である．著者らは，情報処理振興事業協会 (IPA) の独創的情報技術育成事業の支援をうけた「擬人化音声対話エージェント基本ソフトウェアの開発」プロジェクトの中でマルチモーダルコミュニケーションのためのフリーな対話音声合成システムの開発を行っている．本稿では，このシステムの概要，外部インタフェース，合成エンジン，アクセント結合の規則化について紹介する．

2　基本構成

本対話音声合成システムは，独立した四つのモジュール，コマンド解析部，テキスト解析部，音声合成部，音声出力部から成っており，図 1 のような構成をとる．

コマンド解析部は，対話音声合成システムへの入力コマンドを解析し，内部での必要な処理を起動する．テキスト解析部は，発話すべき文テキストを茶筌を用いて解析し，アクセント型を含めた形態素情報を音声合成部へ出力する．音声合成部は，テキスト解析部から渡された形態素解析情報から音声波形を生成

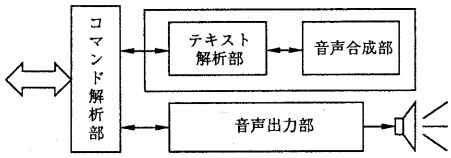

図 1: 音声合成部の構成

```
<SPEECH> <VOICE OPTIONAL="male1">
これは<PRON SYM="アイピーエー">IPA</PRON>の
プロジェクトで開発された<EMPH>対話</EMPH>音
声合成システムです。
</VOICE> </SPEECH>
```

図 2: JEIDA-62-2000 による発話文の記述例

する．音声出力部は，音声波形を出力する．

3 外部インタフェース

3.1 入力コマンド

本システムは，標準入出力を通じたコマンドによって外部と通信する．コマンドは，set, inq, prop の三つで基本的に構成され，それぞれ各種スロット値の設定，問い合わせ，属性変更を行う．例えば，

```
inq SpeakerSet
```

では，利用可能は話者の情報が標準出力に出力される．また，

```
set Text = こんにちは。
```

では,「こんにちは。」の音声が合成され，他モジュールと同期をとために，発話文中の音素系列が継続時間長とともに標準出力に出力される．さらに，

```
set Speak = NOW
```

によって，合成音の出力が直ちに行われる．

3.2 発話文の記述

本システムでは，音声出力する発話文の内容は 3.1 で示した例のように set コマンドで Text スロットの値を設定することによって行う．発話文の表現形式としては，

1. プレインテキストによる漢字仮名混じり文
2. (社) 日本電子工業振興協会「日本語テキスト音声合成用記号の規格 (JEIDA-62-2000)」[2, 3, 4] におけるテキスト埋め込み制御タグおよび仮名レベルの韻律記号に準拠したタグ付きテキスト

を受け付ける．2. の記述例を図 2 に示す．以下，JEIDA-62-2000 の規格のうち，本システムで現在実装されている制御タグについて述べる．

(1) VOICE タグ

```
<VOICE OPTIONAL="話者 ID"> ... </VOICE>
```

は話者 ID を指定する．

```
set Text = <VOICE OPTIONAL="male01"> 彼
女 は 、<VOICE OPTIONAL="female01">は い 。
</VOICE>と言った。</VOICE>
```

のように，一発話内で部分的に別の話者で発話することもできる．

(2) RATE タグ

```
<RATE SPEED="N"> ... </RATE>
```

は，通常の発声に対して N 倍の時間長にする．

(3) VOLUME タグ

```
<VOLUME LEVEL="N"> ... </VOLUME>
```

は，通常の発声に対して音量を N 倍にする．

(4) PITCH タグ

```
<PITCH LEVEL="N"> ... </PITCH>
```

は，通常の発声に対して N 倍のピッチで発話する．また，JEIDA-62-2000 からの拡張として，

```
<PITCH RANGE="N"> ... </PITCH>
```

が，平均ピッチを維持したまま，通常の発声に対するピッチの振れ幅を N 倍に拡大する機能を付加した．

4 合成エンジン

4.1 システムの概要

音声合成エンジンは HMM 音声合成 [5][6] に基づいている．システムは，継続長生成部，基本周波数 (F0) 生成部，パワー生成部，スペクトル生成部，波形生成部から成り，それらは独立したモジュールとして実装されている (図 3)．このため，例えば韻律制御を別個に行ったり，生成された F0 パターンのみを利用するといったカスタマイズが容易に可能である．

システムの入力には，前後の音韻環境，モーラ位置やアクセント型など，音韻，韻律に影響を与える変動要因の組合せ (コンテキスト) を付した音素単位のコンテキストラベル列を与える．システムは入力されたコンテキストラベル列に対応するコンテキスト依存モデル (HMM) を連結して一発話単位の HMM を構成し，この HMM から音声合成に必要なパラメータを生成する．

HMM 音声合成では，音素を単位として，スペクトル，パワー，F0，継続長がコンテキスト依存モデルにより HMM の枠組で統一的にモデル化されている [6]．モデル構築の際，可能な全てのコンテキストの組合せを網羅する学習データを用意することは現実的には不可能であるため，コンテキスト依存モデルは決定木 (二分木) に基づくコンテキストクラスタリングがされている．これにより，学習データに出現しないコンテキストの組合せに対しても対応するモデルが一意に決定される．実際には，このコンテキストクラスタリングをスペクトルとパワー，ピッチ，継続長それぞれに適用し，別々に決定木を作成している．

本システムでは，男女各 1 名の話者について，ATR 音韻バランス文 (503 文章) 及び日本音響学会研究用音声データベース案内タスク文の一部 (447 文章) から構築したコンテキスト依存モデルが用意されている．

4.2 継続長生成部

各音素 (実際には音素を単位とするコンテキスト依存モデル，以下同様) の継続長は多次元ガウス分布でモデル化されており，多次元ガウス分布の各次元は HMM の各状態の継続長分布を表している [6]．

モデルの継続長が N 次元のガウス分布でモデル化されているとして，音素列 $(p_1, p_2, \ldots, p_L)$ に対して全体の発話時間 T が与えられた場合，各音素 p_l の継続長は次式で与えられる [5]．

$$d_{p_l} = \sum_{n=1}^{N} \mu_{p_l n} + \rho \sum_{n=1}^{N} \sigma^2_{p_l n} \qquad (1)$$

ただし，$\mu_{p_l n}$, $\sigma^2_{p_l n}$ はそれぞれ音素 p_l の継続長分布の各次元 n $(1 \leq n \leq N)$ の平均と分散であり

$$\rho = \left(T - \sum_{l=1}^{L} \sum_{n=1}^{N} \mu_{p_l n} \right) \Big/ \sum_{l=1}^{L} \sum_{n=1}^{N} \sigma^2_{p_l n} \qquad (2)$$

とする．なお，全体の発話時間 T が指定されない場合には $\rho = 0$，すなわち状態継続長分布の平均値の和として決定される．

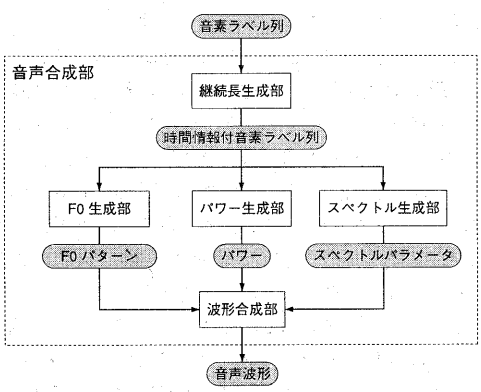

図 3: 音声合成部

4.3 基本周波数 (F0) 生成部

F0 生成部では，時間情報 (音素継続長の情報) 付コンテキストラベル列を入力として基本周波数パターンを生成する．各音素の F0 パターンは，多空間上の確率分布に基づく隠れマルコフモデル (MSD-HMM) [7][8] でモデル化されており，MSD-HMM の各状態は，有声となる確率，無声となる確率，有声である場合の対数基本周波数およびその動的特徴量の平均と分散，および単一ガウス分布で近似した状態継続長分布を持つ．

まず，与えられた各音素の音素継続長に対して，モデルの状態継続長分布に基づいて式 (1)，(2) により各状態の継続長を決定する．有声となる確率が無声となる確率よりも大きい状態を有声の状態，小さい状態を無声の状態とし，有声の状態が継続する区間を有声区間，無声の状態が継続する区間を無声区間とする．そして，有声区間内で尤度最大化基準に基づくパラメータ生成手法 [9] に基づいてピッチパターンを生成する．

4.4 スペクトル生成部とパワー生成部

スペクトル生成部及びパワー生成部では，時間情報 (音素継続長の情報) 付音素ラベル列を入力としてメルケプストラムベクトル列を生成する．

各音素のスペクトル列は，メルケプストラムをパラメータとして通常の連続分布 HMM でモデル化されている．スペクトル生成時には，F0 パターン生成と同様，各状態の継続長を決定し，次に尤度最大化基準に基づくパラメータ生成手法 [9] に基づいてメルケプストラム列を生成する．

パワーはスペクトルパラメータとして用いられているメルケプストラムの 0 次項として，スペクトルパラメータとともに連続分布 HMM でモデル化されており，スペクトル生成と同様，尤度最大化基準に基づくパラメータ生成手法に基づいてメルケプストラムを生成し，その 0 次項の値をパワーに変換する．

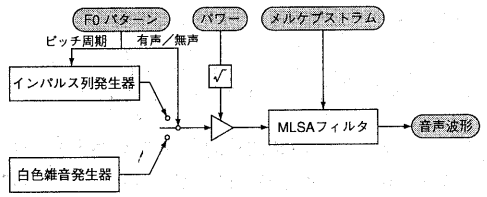

図 4: 波形生成部

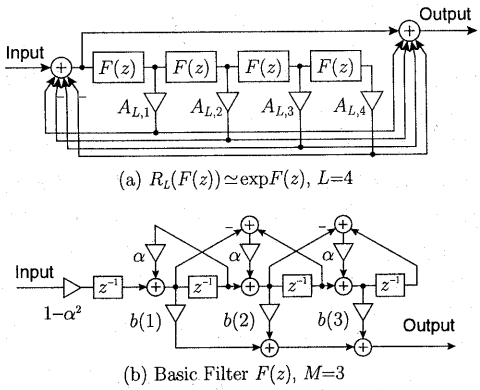

図 5: MLSA フィルタ

4.5 波形生成部

波形生成部では，F0 生成部，パワー生成部，スペクトル生成部で生成された F0 パターン，パワー，メルケプストラム列を入力とし，ボコーダ方式により音声波形を生成する (図 4)．入力された F0 パターンから有声／無声およびピッチ周期に従って音源信号を生成し，パワーを掛け合わせ，メルケプストラムをパラメータとする MLSA フィルタ [10][11] を励振することにより，その出力として音声波形が得られる．

本システムで使用している指数関数の有理式近似に基づく MLSA フィルタの構成を図 5 に示す．ここで，基礎フィルタ $F(z)$ の係数 $b(m)$ は，メルケプストラム係数を $c(m)$ として

$$b(m) = \begin{cases} c(m), & m = M \\ c(m) - \alpha b(m+1), & 0 \leq m < M \end{cases} \quad (3)$$

で与えられる．ここで M はスペクトルモデルの次数である．本システムでは，サンプリング周波数 16kHz，メルケプストラム次数 $M = 24$ であり，周波数伸縮パラメータ α の値はメルスケールを近似する $\alpha = 0.42$ としている．さらに，有理近似式の係数 $A_{4,l}$ には文献 [10] で示されている値を用いている．

5 アクセント結合の規則化とデータベース化

5.1 テキスト音声合成が必要とする韻律情報

テキスト (漢字仮名混じり文) を入力として音声を合成する場合，入力文に対して言語解析を施し，音韻情報と韻律情報を抽出する必要がある．後者の情報としては，呼気段落境界位置，アクセント句境界位置，及び，アクセント句内におけるアクセント核位置などが相当する．周知のように日本語では，各単語毎に固有の単語アクセントが定義されるが，これら単語が文中にて連接した場合アクセント型が変形することが多い (アクセント結合)．本節では，「入力テキストに対してアクセント句境界が与えられた場合に，そのアクセント句内にアクセント核を置く必要が有るのか無いのか，前者の場合には，どの位置に置く必要があるのか」に関する規則化，及びその規則を具体的に構築するために必要なデータ収集に関して述べる．

5.2 アクセント結合の規則化

匂坂らはテキスト音声合成を念頭に置いて，日本語におけるアクセント結合の現象を網羅的に説明する規則を構築している．この規則はテキスト音声合成の分野において広く利用されており，本節でも以降，この規則をベースとして議論する．アクセント句は 1 つ以上の文節から構成されるが，匂坂らによれば，文節アクセントパターンの結合は比較的簡単な規則で記述可能であり，その結果，規則化の焦点は文節を構成する場合のアクセント核位置となる．本報告では紙面の都合上，文節＝自立語＋付属語，におけるアクセント結合の様子を例にとって議論する．アクセント変形が観測される，自立語＋付属語以外の語連鎖である複合単語，接尾語連接，接頭語連接における規則化も同様の検討を進めることで可能となる．

自立語 (名詞，動詞，形容詞) に付属語 (助詞，助動詞) が結合して文節を作る場合，文節のアクセント型は，自立語のモーラ長，アクセント型，及び付属語のモーラ長，アクセント結合属性 (表のアクセント価及びアクセント結合様式) によって説明される．表 1 に付属語連接におけるアクセント結合の様子を示す．ここでアクセント価 $\tilde{M}_2$ とは，連接時に，付属語の先頭から第 $\tilde{M}_2$ モーラ前にアクセント核が付与される

表 1: 自立語＋付属語による文節のアクセント型
N_1 モーラ M_1 アクセント型 [自立語] ＋
N_2 モーラ $\tilde{M}_2$ アクセント価 [付属語]
→ N_c モーラ M_c アクセント型 [文節]

アクセント結合様式	文節アクセント M_c	
	$M_1 = 0$ の場合	$M_1 \neq 0$ の場合
F1 (従属型)	$M_1(=0)$ 笑うほど	M_1 歩┐いた
F2 (不完全支配型)	$N_1 + \tilde{M}_2$ 笑った┐り	M_1 歩┐くヨウダ
F3 (融合型)	$M_1(=0)$ 笑いながら	$N_1 + \tilde{M}_2$ 歩かせ┐る
F4 (支配型)	$N_1 + \tilde{M}_2$ 笑うま┐い	歩きま┐す
F5 (平板化型)	0 笑うだけ	歩くだけ

表 2: 付属語アクセント結合属性決定手順

■ 手順 1
「付属語」に「歩く (有核)」を接続する.
if 「歩く＋付属語」が, 0 型 (無核) であれば, その付属語の属性を F5 と定義する.
else 「歩く＋付属語」が, 2 型 (「歩く」の核位置) であれば, 手順 2 へ
else 「歩く＋付属語」が, 3(「歩く」のモーラ数) ＋ N 型であれば, 手順 3 へ
■ 手順 2
「付属語」に「笑う (無核)」を接続する.
if 「笑う＋付属語」が, 0 型 (無核) であれば, その付属語の属性を F1 と定義する.
else 「笑う＋付属語」が, 3+N 型 (有核) であれば, その付属語の属性を F2/N とする (N は整数).
■ 手順 3
「付属語」に「笑う (無核)」を接続する.
if 「笑う＋付属語」が, 0 型であれば, その付属語の属性を F3/N とする (N は手順 1 より).
else 「笑う＋付属語」は, 3+N 型のはずなので (N は手順 1 より), その付属語の属性を F4/N とする.

ことを意味する. なお, $\tilde{M}_2$ は付属語を単独で発声した場合のアクセント型とは必ずしも一致しない.

5.3 付属語のアクセント結合属性の推定

さて, 上記規則を音声合成に適用する場合は, 全ての付属語に対してアクセント結合属性 (アクセント価及びアクセント結合様式) を定義する必要がある. この場合, 各付属語に対して**表 1** における全てのアクセント結合属性を検討する必要は無く, 例えば**表 2** に示す手順を踏むことで作業の効率化が図れる. 昨年度の IPA の活動では, 東京方言話者と考えられる話者 10 人に対して, この手順をそのまま実行させ, 各付属語のアクセント結合属性の推定作業を行なった.

なお実験に先だって, 無意味モーラ列に対するアクセント型同定を行なわせ, 各被験 i の正答率 $w_i(w_i = 0 \sim 1.0)$ を算出した. この値を「各被験者の信頼性」として利用し, 最終的な集計作業に用いた. また, ある単語に対して異なる被験者が異なる型を想定する場合があるが, これはアクセントが本来持つ「揺れ」として捉えた.

上記の実験の結果, 以下の様な形でアクセント属性が推定されることとなる.

例:
(品詞 (助動詞)) ((見出し語 (た 0)) (読み タ) ....)
F1(0.7875) F2/-1(0.2947) (0.0000)
F2/1(0.1353) F3/0(0.0773) (0.0000)
F3/0(0.0773) (0.0000) (0.0000)

動詞接続時 形容詞接続時 名詞接続時

F2/1 はアクセント属性が F2, アクセント価が 1 であることを意味する. アクセント価を要しない属性には記されていない. 括弧内の数字は, 各アクセント属性に対する信頼度であり, $S = \sum w_i$ とした場合, [そのアクセント属性を採用した被験者の信頼度の和 $]/S$, として計算される. 信頼度の総和が 1.0 にならないのは, 推定結果として「推定不能」が選択されたことによる. 図 6 は, 昨年度に行なった推定実験結果より算出した, 助動詞のアクセント属性の揺れの様子である. 即ち, 信頼度が θ 以上の属性を持つ語が全体の何割がを示している. これを見ると, 属性信頼度が 0.6 以上の助動詞は, 約半数しか無いことが分かる[1].

昨年度の推定実験に対する反省事項として, 1) 推定手順の複雑さ, あるいは, 被験者の日本語文法知識不備[2]により, 推定手順遂行が信頼性高く行なわれなかった可能性がある, 2) 被験者数, 推定作業の際に用いられた (自立) 語数の増加, などが挙げられていた. これらの解決を目的として, 約 1,800 語の付属語, 接頭語, 接尾語に対して現在推定再実験を遂行中である. 被験者タスクとしては,「与えられた語系列 (自立語＋付属語, 接頭語＋自立語, 自立語＋接尾語など) 中のアクセント核位置を答えさせる」という形態とした. この場合, 音声による確認作業が必要であると考え, 選択したアクセント核位置の合成音を聞かせ, 逐一確認作業を行なわせることとした. 図 7 がデータ収集用のインターフェイスである. なお, 文節のテキスト提示のみでは意味が唯一に決定できな

[1] 但し,「推定不能」も揺れの一部として算出されている.
[2] 推定結果例における, 形容詞接続時の信頼度総和が 1.0 にならないのは, 形容詞に助動詞「た」を連接させた場合に, どこまでが形容詞の語幹なのか分からない学生が多かったことを意味する.

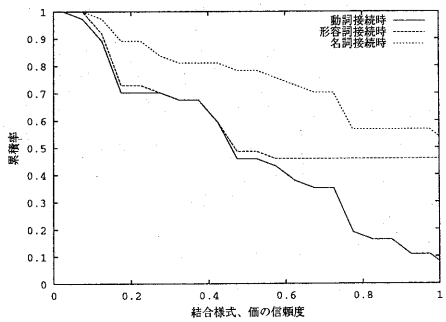

図 6: 助動詞に対するアクセント属性の揺れ

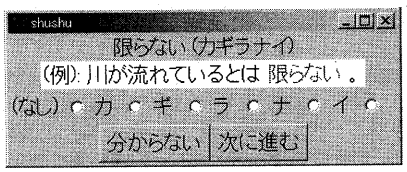

図 7: アクセント結合属性推定実験

い場合があり，実際の文/句例も示している．得られたアクセント核位置結果に対して，**表 2** にあるようなアクセント結合属性決定手順を踏むことで種々の属性値は決定可能であるが，収集データを，学習データ／評価データに区分することで，規則の評価，改良についても検討する予定である．

なお，上述した文節内のアクセント核位置決定規則を運用する場合，巡回的適用則，音節内移動規則，無声化に伴う移動規則など，など種々の事項に注意する必要があるが，これについては参考文献を参照されたし．

6 まとめ

現在開発を進めているマルチモーダルコミュニケーションのための対話音声合成システムについて述べた．今後は，アクセント型辞書の整備と茶筌への組み込み，音声出力の中断機能，などの実装を進めていく予定である．

謝辞

本研究は，情報処理振興事業協会 (IPA) の支援を受けた「擬人化対話エージェント基本ソフトウェアの開発」プロジェクトの一部として行われている．プロジェクト参加者・関係者および情報処理振興事業協会に感謝する．

参考文献

[1] 山下洋一：“対話システムにおける音声合成”, 情報処理学会研究報告, SLP-33-4, pp.19-24 (2000).

[2] 赤羽誠, 養輪利光, 板橋秀一：“音声合成用記号の標準化について”, 音響誌, 57, 12, pp.776-782 (2001).

[3] 養輪利光, 赤羽誠, 板橋秀一：“JEIDA 日本語テキスト音声合成用記号”, 日本音響学会秋季講演論文集, 2-1-5, pp.183-184 (2000).

[4] (社) 日本電子工業振興協会：日本語テキスト音声合成用記号の規格, JEIDA-62-2000 (2000).

[5] 益子 貴史, 徳田 恵一, 小林 隆夫, 今井 聖, “動的特徴を用いた HMM に基づく音声合成”, 信学論 (D-II), **J79-D-II**, 12, pp.2184–2190, Dec. 1996.

[6] 吉村 貴克, 徳田 恵一, 益子 貴史, 小林 隆夫, 北村 正, “HMM に基づく音声合成におけるスペクトル・ピッチ・継続長の同時モデル化”, 信学論 (D-II), **J83-D-II**, 11, pp.2099–2107, Nov. 2000.

[7] 徳田 恵一, 益子 貴史, 宮崎 昇, 小林 隆夫, “多空間上の確率分布に基づいた HMM”, 信学論 (D-II), **J83-D-II**, 7, pp.1579–1589, Jul. 2000.

[8] 益子 貴史, 徳田 恵一, 宮崎 昇, 小林 隆夫, “多空間確率分布 HMM によるピッチパタン生成”, 信学論 (D-II), **J83-D-II**, 7, pp.1600–1609, Jul. 2000.

[9] 徳田 恵一, 益子 貴史, 小林 隆夫, 今井 聖, “動的特徴を用いた HMM からの音声パラメータ生成アルゴリズム”, 日本音響学会誌, **53**, 3, pp.192–200, Mar. 1997.

[10] T. Fukada, K. Tokuda, T. Kobayashi and S. Imai, “An Adaptive Algorithm for Mel-Cepstral Analysis of Speech,” Proc. ICASSP92, pp.137–140, 1992.

[11] 今井 聖, 住田 一男, 古市 千枝子, “音声合成のためのメル対数スペクトル近似 (MLSA) フィルタ”, 信学論 (A), vol.J66-A, no.2, pp.122–129, Feb. 1983.

[12] NHK 編, “日本語発音アクセント辞典 (改訂新版)”, 日本放送出版協会 (1985)

[13] 金田一春彦監修, 秋永一枝編, “明解日本語アクセント辞典 (第二版)”, 三省堂 (1981)

[14] 平山輝男, “全国アクセント辞典”, 東京堂出版 (1955)

[15] 天野成昭, 近藤公久編著, “日本語の語彙特性 (単語アクセント)”, 三省堂 (1999)

[16] 匂坂芳典, 佐藤大和, “日本語単語連鎖のアクセント規則”, 信学論, Vol.J66-D, No.7, pp.849–856 (1983)

[17] 佐藤大和, “複合語におけるアクセント規則と連濁規則”, 日本語と日本語教育, 第二巻, 明治書院 (1989)

[18] 宮崎正弘, “単語間の意味的結合関係を用いた複合語アクセント句の自動抽出法”, 信学論, Vol.J68-D, No.1, pp.25–32 (1985)

[19] 橋本新一郎, “日本語単語アクセントの緒性質”, 信学論, Vol.J56-D, No.11, pp.654–661 (1973)

[20] 桜井淳宏, 芦田直之, 広瀬啓吉, “音声データベースのための複合名詞におけるアクセント変形タイプの自動推定”, 日本音響学会春季講演論文集, 3-3-6, pp.255–256 (1999)

[21] 森田真弘, 瀬戸重宣, 籠嶋岳彦, 赤嶺政己, “モーラを単位としたアクセント規則の自動構築”, 日本音響学会秋季講演論文集, 1-2-19, pp.211–212 (1998)

[22] 蜂谷雅弘, 森山高明, 小川均, 天白成一, 橋本雅行, “アクセント法則を取り入れたアクセント型推定手法”, 日本音響学会秋季講演論文集, 2-2-9, pp.241–242 (1997)