音声駆動型身体引き込みキャラクタ InterActor が 発話音声に与える効果*

檀 原 龍 正^{*1}, 渡 辺 富 夫^{*2}, 大久保 雅史^{*3}

Effects of Speech-Driven Embodied Entrainment Character “InterActor” on Talker’s Speech

Ryusei DANBARA^{*4}, Tomio WATANABE and Masashi OKUBO

^{*4} Graduate School of System Engineering, Okayama Prefectural University,
111 Kuboki, Soja-shi, Okayama, 719-1197 Japan

A speech-driven embodied entrainment character called “InterActor” with both functions of speaker and listener has been developed for activating human interaction by generating the whole body motions such as nodding, blinking and the actions of head, arms and waist coherently related to speech input. This paper demonstrates the effects of InterActor on interaction in human speech input by the analysis of speech temporal parameters. It is found that speech duration increases and speech activity decreases because of the increase of total silence duration and the number of silence parts in human-to-InterActor speech input, which is the advantage of real time processing for speech recognition and so forth. Interaction with InterActor is also highly evaluated in smoothness, enjoyment and preference by sensory evaluation.

Key Words: Human Communication, Human Interaction, Embodied Interaction, Interaction Sharing

1. はじめに

対面コミュニケーションでは単に言葉だけでなく、発話音声に基づいて頷きや身振り手振りなど、身体リズムを共有し、相互に引き込むことでインタラクションしている。この身体的引き込みが対話時の一体感を生み、対話相手とのかわりを実感させている⁽¹⁾。情報機器を介したコミュニケーションにおいても、このメカニズムを導入することで、インタラクションを円滑にし、一体感のあるコミュニケーションを支援できるものと期待される。

近年、情報機器の進展とともに、遠隔地の人とのコミュニケーションや情報機器とのコミュニケーションを支援する自然なインタフェースとして、コミュニケーションにおける身体性に着目したヒューマンエージェントやアバタを用いた研究がさかに行われている^{(2) - (5)}。しかしながら、対面でのコミュニケーションにおける頷きや身振りなどの身体リズムの引き込

みに基づいたモデルを CG やロボットなどのメディアに導入することで、一体感が実感できるコミュニケーション支援システムの研究開発は少ない⁽⁶⁾。

著者らは既に、身体的インタラクションロボット InterRobot を用いた身体的コミュニケーション支援システムのプロトタイプを開発している⁽⁷⁾。さらに高度メディア社会での広範囲な応用を目指して、InterRobot の電子メディア版として、聞き手と話し手の両機能を有し、豊かなコミュニケーション動作を音声のみから自動生成する CG キャラクタである音声駆動型身体引き込みキャラクタ InterActor を開発し、官能評価によりコミュニケーション支援に有効であることを示している (図 1)^{(8) - (10)}。

Fig.1 Remote communication by using InterActor.

* 原稿受付 2004 年 11 月 26 日。

^{*1} 正員, 岡山県立大学大学院情報系工学研究科(〒719-1197 総社市窪木 111)。

^{*2} 正員, 岡山県立大学情報工学部(〒719-1197 総社市窪木 111), 科学技術振興機構 CREST。

^{*3} 正員, 同志社大学工学部(〒610-0321 京田辺市多々羅都谷 1-3)。

E-mail: ryusei@eco.cse.oka-pu.ac.jp

InterActor は身体リズムの引き込みの重要性に着目したアバタであり、音声と身体動作の関係に着目し、モデル化することで発話音声からコミュニケーション動作を自動生成し、身体リズムの引き込みによりインタラク션을円滑にし、コミュニケーションを支援する。InterActor への入力情報は音声のみで、InterActor を介したコミュニケーションは自己の映像情報の匿名性を活かした新しいモードのコミュニケーションである。

InterActor の音声に基づく身体反応動作は音声入力の ON-OFF のみに対応させたランダム動作と比較して有意にコミュニケーション支援に有効であることが官能評価によって確認されている⁽⁹⁾。これは、単に InterActor が動くから支持されたのではなく、InterActor の発話音声に基づく引き込み動作の有効性を示している。発話者は InterActor の反応を見ながら発話し、InterActor はその発話音声に基づいて身体動作を生成する。これを相互に繰り返すことでインタラクションしている。そのため、InterActor の反応が発話者自身の発話特性に影響を与えていると考えられる。

本論文では、聞き手としての InterActor に語りかけたときの話し手の音声、とくに発話音声の時間パラメータに着目し、InterActor が話し手の発話特性に与える影響を定量的に評価している。具体的には、InterActor の反応条件として、全く動かない条件（反応がない場合）と発話音声に基づいて引き込み動作を行う条件（反応がある場合）を設定し、各条件での話し手の発話音声を計測して比較検討することで、InterActor が話し手の発話音声に与える効果を明らかにしている。

2. InterActor

2.1 システムの概要

InterActor の概略を図 2 に示す。システムの構成には Microsoft Windows2000/XP を OS としてサウンドカードを搭載する PC、マイクなどの音声入力装置、スピーカなどの音声出力装置を必要とする。プログラムのソースコードは Microsoft Visual C++ Ver6.0 により記述した。InterActor は対話者の代役（アバタ）としてディスプレイ上に表示される電子メディアの CG キャラクタであり、その描画には Microsoft DirectX 6.0 のコンポーネントを利用している。この InterActor は頷き、瞬き、口の開閉に加え、各関節部位（首・腰・肩・肘・手首）の合計 19 自由度の曲げ動作、回転動作を組み合わせることで、話し手及び聞き手としての多様なコミュニケーション動作を表現することができる。

InterActor を用いた身体的インタラクションシステムのコンセプトを図 3 に示す。図中の対話者 1 が InterActor 2 に語りかけると、InterActor 2 が聞き手として頷き・瞬きや身振りなど身体全体で反応する。同時に、対話者 1 の音声ネットワーク等を経由して対話相手方の InterActor 1 に送信される。InterActor 1 は、その音声時系列に基づいて話し手としてコミュニケーション動作をその音声に関連付けて生成し、音声と動作を同時に提示することで、発話情報を効果的に対話者 2 に伝達することができる。この操作を相互に連続して行うことで、円滑なコミュニケーションが実現される。

Fig.2 Outline of InterActor.

Fig.3 Concept of embodied interaction system.

2.2 InterActor のインタラクションモデル

対面コミュニケーションでの音声と身体動作の関係を分析した結果、話し手の発話音声と聞き手の身体動作では、頷きを主体とする頭部の動きとの相関が最も高いこと、さらに胴部・腕部の動作は話し手の発話音声よりも聞き手自身の頭部の動作との関係が強いこ

Fig.4 Listener's interaction model.

とがわかった⁽⁹⁾。その解析結果に基づき、聞き手のインタラクションモデルとして、音声の ON-OFF パターンに基づく頷き反応モデルと、腕部および上部に対して頷きの予測値に基づく身体動作モデルを導入している⁽¹¹⁾。聞き手の動作の予測モデルを図 4 に示す。頷きの予測モデルはマクロ層とミクロ層からなる階層モデルである。マクロ層では音声の呼気段落区分での ON-OFF 区間からなるユニット区間に頷きの開始が存在するかを予測する。予測には $[i-1]$ ユニット以前のユニット時間率 $R(i)$ (ユニット区間での ON 区間の占める割合、(2) 式) の線形結合で表される (1) 式の MA (Moving-Average) モデルを用いる。雑音は身体のゆらぎを示すノイズ成分である。

$$M_u(i) = \sum_{j=1}^J a(j)R(i-j) + u(i) \quad (1)$$

$$R(i) = \frac{T(i)}{T(i) + S(i)} \quad (2)$$

$a(j)$: 予測係数

$T(i)$: i 番目ユニットでの ON 区間

$S(i)$: i 番目ユニットでの OFF 区間

$u(i)$: 雑音

$$M(i) = \sum_{j=1}^K b(j)V(i-j) + w(i) \quad (3)$$

$b(j)$: 予測係数

$V(i)$: 音声データ

$w(i)$: 雑音

予測値 $Mu(i)$ が閾値 0.20 を越えて、頷きが存在すると予測された場合には、処理はミクロ層に移る。ミクロ層では音声の ON-OFF データ (30Hz, 60 個) を入力とし、(3) 式を用いて MA モデルで頷きの開始

時点を推定する。予測値が閾値 0.46 を越えた場合には InterActor を頷かせる。瞬きについては、対面コミュニケーション時における瞬き特性に基づいて、頷きと同時に瞬きさせ、それを基点として指数分布させている⁽¹²⁾。身体反応の推定には頷きの予測値を用い、頷きよりも低い閾値で InterActor の各部位 (頭部、胸部、右肘、左肘) のうちいくつかを選択して動作させることで頷きと関係付けている。

話し手のモデルについても同様に、対面コミュニケーション時の音声と身体動作の特性から、音声の ON-OFF パターンに基づく身体全体の動作を予測するモデルと音声の振幅に基づく腕部動作モデルを導入している。身体動作モデルとしては全ての動きの ON-OFF の総和データから体の動くタイミングを予測し、閾値を越えたときに InterActor の各部位 (頭部、胸部、右肘、左肘) のうちいくつかを選択し動作させることで発話音声と関係付けている。

したがって、InterActor は言葉の意味を理解して反応するわけではなく、InterActor の音声に基づく身体動作の引き込み反応に対してユーザが意味付けを行い、その意味付けに従ってユーザは自分自身の振る舞いを変化させる。この InterActor の身体動作とユーザの意味付けが相互に繰り返されることでインタラクションが行われる。

InterActor は単調な身体動作による反応よりも、多様な身体動作による反応を行った方が、この意味付けの幅が広くなり、より効果的なインタラクションが行われることが確認されている⁽¹³⁾。また、システムの入力情報は音声のみであり、マイクからの実時間入力ばかりでなく、あらかじめ記録された音声データを用いることができる。このため、録音された過去の音声情報から身体動作を生成することで、時間を超えたコミュニケーションでの身体性の共有も可能である。

3. 発話実験

3.1 実験概要

音声に基づく InterActor の身体動作が、話し手の音声そのものに与える効果を検証するために発話実験を行った。実験においては、留守番電話にメッセージを録音する場面を想定し、実際には被験者に表 1 に示すメッセージを InterActor に対し語りかけさせ、その発話音声を記録した⁽¹⁴⁾。

Table 1 Message used in the experiment.

もしもし、岡山県立大学の〇〇です。
本日の論文の打合せですが、
明日の3時に変更になりました。
よろしくお願いします。
*〇〇部分は被験者名

発話実験においては、InterActor の状態として、InterActor を表示しているだけで反応がない場合と、聞き手としての引き込み動作反応がある場合の2種類の条件を設定した。図 5 に InterActor のコミュニケーション動作の一例を示す。

Fig.5 Example of communication motion of InterActor.

実験手順は、まず反応がない場合で行い、続けて反応がある場合で行った。これを 1 セットとして、1 人の被験者に対して 3 セット繰り返し行った。この際、全てのセットを連続して行うことで順序効果の可能性を排除している。また、これまでの研究成果から音声に関連付いたコミュニケーション動作がランダム動作に比べて有意に好まれていることが既に確認されているので⁽⁹⁾、本実験では InterActor によるランダム動作を実験条件に含んでいない。

実験終了後、反応条件ごとに発話者の音声情報を解析・比較することで、InterActor の身体動作の有無による話し手の音声への影響を定量的に評価した。被験者は学生 26 名である。

実験システムの構成を図 6 に示す。マイクから PC (CPU: Athlon-1GHz) へ入力された音声は、サンプリング周波数 11,025Hz で A/D 変換され、リアルタイムで InterActor の動作生成信号として用いられるとともに、

に、DAT (TEAC RD-130TE) に録音された。実験の様子は CCD カメラによって撮影して DV レコーダで収録した。記録された発話実験の様子の一例を図 7 に示す。

Fig.6 Experimental setup.

Fig.7 Example of experimental scene.

3.2 InterActor の身体動作

発話実験における発話者の入力音声と、その音声に基づく InterActor の身体動作 (頭部、胴部) を相互相関関数 $C(\tau)$ で分析した。

$$C(\tau) = \frac{\sum_{i=1}^{n-\tau} \{x(i) - \mu_x\} \{y(i+\tau) - \mu_y\}}{\sqrt{\sum_{i=1}^n \{x(i) - \mu_x\}^2} \sqrt{\sum_{i=1}^n \{y(i) - \mu_y\}^2}} \quad (4)$$

μ_x : 音声 $x(i)$ の平均値

μ_y : 傾き $y(i)$ の平均値

n : データ数

τ : 時間遅れ

音声情報は、サンプリング周波数 30Hz で平均入力音圧が閾値 (発話しない状態での平均入力音圧に 12dB を加えた値) を超えた場合を ON (1)、そうでない場合を OFF (0) とし、ON 区間に対して 5/30 s のハングオーバー (ON 区間を伸ばして短い OFF 区間を削除) 処理を施したものを音声の ON-OFF データとして使用した。InterActor の動作にはディスプレイに描画する際の制御信号を用い、MA モデルによる予測値に基づき、身体各部について動作中を ON 停止中を OFF とした。聞き手としての InterActor の身体

動作の分析結果の一例を図 8 に示す。(a) は入力音声とその ON-OFF データ、さらに、その ON-OFF データから MA モデルにより推定した InterActor の頷きの予測値と予測値に基づいて生成された InterActor の各身体動作のデータである。胸部動作の生成が、頭部動作の生成に先行しており、頻度も高くなっていることがわかる。(b) は入力音声に基づく InterActor の各身体動作の相互相関関数 $C(\tau)$ の分析結果である。

Fig.8 Time series of talker's speech and the motions of InterActor as listener, and cross-correlation $C(\tau)$ between them.

この結果から、入力音声に対する頷きの相互相関関数 (Speech-Head) は約 0.1s に負の相関を示し、胸部との相互相関関数 (Speech-Body) は約 0.2s に負の相関を示していることがわかる。これは、頷きよりも少し早い時点で身体が動き始めるという、対面コミュニケーション時の人の動作特性に合致している⁽¹⁵⁾。つまり、身体動作の反応には頷きの予測値 0.46 よりも低い閾値 0.20 で各部位を動作させることで、頷きを中心に身体全体の動きを伴う自然なコミュニケーション動作の実現に繋がっている。また、全体的な動き (Speech-Head and Body) としては約 0.1s に負の相関を示していることから、音声 ON 区間の終了を予測して動作していることがわかる。全被験者 26 名の入力音声に対する頷きの相互相関関数の最小値の平均値 (標準偏差) は、-0.55 (0.1)、胸部とでは -0.24 (0.06)、身体動作全体では 0.36 (0.12) であった。

3.3 官能評価

すべての実験終了後に被験者に対して、InterActor の反応がない場合を基準に、反応がある場合について「話し易さ」や「楽しさ」などを 7 段階 (中立 0) で評価させたアンケート結果を図 9 に示す。どの項目に

ついても肯定的な評価が得られ、語りかけに対する聞き手としての InterActor の有効性が示されている。とくに「話し易さ」については、ほとんどの被験者が肯定的な回答をし、さらにアンケートの自由記述欄においても、InterActor が頷くことで話しのタイミングが掴み易かったという回答をした被験者もいた。

Fig.9 Result of questionnaire.

3.4 発話音声の解析

官能検査において InterActor の聞き手としての効果が表示された理由を検証するために、発話実験において、InterActor の反応が発話音声の時間パラメータに与える影響について検討する。まず、InterActor の反応がない場合とある場合の話し手が話し始めてから話し終わるまでの発話時間を比較した。図 10 に被験者の発話時間の平均値と標準偏差をセットごとに示す。この図より、全てのセットにおいて InterActor の反応がない場合に比べて反応がある場合で発話時間が長くなっていることがわかる。セットごとに対応関係のある平均値の差の検定をした結果、全てのセットで有意水準 1% で有意差が認められた。

Fig.10 Influence of InterActor's motion on the speech time.

つぎに、発話区間内における ON 区間 (有音区間) のみを抽出した全 ON 区間の平均値と標準偏差を

図 11 に、OFF 区間（無音区間）のみを抽出した全 OFF 区間の平均値と標準偏差を図 12 に示す。

全 ON 区間に関しては、InterActor の反応の有無による顕著な差は見られない。一方、全 OFF 区間では反応がある場合に全 OFF 区間が長く、全てのセットで有意水準 1% で有意差が認められた。これらのことより、InterActor の反応がある場合は、反応がない場合に比べ、全 OFF 区間が長くなることで全体として発話時間が延びていることが明らかとなった。

Fig.11 Total ON duration.

Fig.12 Total OFF duration.

また、音声情報処理の分野でよく用いられる時間率について検討する。時間率は、発話区間内での ON 区間の占める割合を示す代表的な発話の特徴パラメータであり、(5) 式によって表わされる⁽¹⁶⁾。

$$R = \frac{\bar{T}}{\bar{T} + \bar{S}} \quad (5)$$

R : 時間率

$\bar{T}$: 発話音声の平均 ON 区間

$\bar{S}$: 発話音声の平均 OFF 区間

図 13 に発話実験における被験者の時間率の平均値と標準偏差を示す。全てのセットにおいて、反応がある場合とない場合の間に有意水準 1% で有意差があり、反応がある場合に時間率の低下が認められた。

Fig.13 Speech activity.

また、(5) 式に用いられている各被験者の平均 ON 区間の平均値と標準偏差を図 14 に、平均 OFF 区間の平均値と標準偏差を図 15 に示す。

Fig.14 Average time of ON duration.

図 14 より、InterActor の反応がない場合に比べ、反応がある場合で平均 ON 区間が短くなっている。全てのセットで、有意水準 1% で有意差が認められた。一方、図 15 より平均 OFF 区間では反応の有無による顕著な差は見られない。

Fig.15 Average time of OFF duration.

つぎに、発話区間における OFF 区間数の平均値と標準偏差を図 16 に示す。全てのセットにおいて有意水準 1% で有意差が認められ、反応がある場合に OFF 区間の数が増えていることがわかる。

Fig.16 Number of OFF parts.

これらの結果より、図 11 で示した全 ON 区間に関しては反応の有無による差はなく、図 13 に示した平均 ON 区間に関しては、反応がある場合で短くなることは、InterActor が話し手の音声に対して反応することにより、反応がない場合には見られなかった新たな OFF 区間が生じるためであると考えられる。新たな OFF 区間が生じることで、反応がない場合には 1 つであった ON 区間が、反応がある場合では 2 つの ON 区間に分割されることになる (図 17)。つまり、全体として ON 区間の数が増加する。ON 区間数の増加は OFF 区間数の増加と等しく、その結果は図 16 に示した結果と合致する。

Fig.17 Division of ON duration caused by the effects of InterActor.

InterActor の反応がない場合とある場合について、ある被験者のメッセージの同じ箇所の音声と、それに対応する音声の ON-OFF データの一例を図 18 に示す。

(a) の反応がない場合に比べ (b) の反応がある場合の方が OFF 区間が多くなっていることがわかる。つぎに、文と文の切れ目であり、どの被験者も明白に間を取る代表的な OFF 区間に着目し、音声の時間パラメータの変化を比較した。具体的な OFF 区間として、「もしもし岡山県立大学の〇〇です」と「本日の論文の打ち合わせ…」との間に存在する OFF 区間に着目した。この OFF 区間の平均値と標準偏差を図

Fig.18 Influence of InterActor on the talker's speech.

19 に示す。1 セット目については、有意水準 1% で有意差が認められ、反応がある場合で OFF 区間が長くなっている。2, 3 セット目については差があるものの、有意差は認められなかった。

上の結果より、発話時間が延びる要因として呼吸段落区分での OFF 区間が大きく影響しており、OFF 区間数が増えることによって発話時間が延びていることがわかる。同時に、明白に間を取る OFF 区間については InterActor が反応することで OFF 区間が長くなっていることがわかる。これらのことは、発話者がインタラクションを確認しながら InterActor に語りかけるためであると考えられる。

Fig.19 Representative OFF duration.

4. まとめ

音声駆動型身体引き込みキャラクタ InterActor を用いた発話実験を行い、官能評価を行うとともに、InterActor の反応が話し手の発話特性に与える効果を定量的に解析した結果、以下に示す結論を得た。

- (1) 官能評価により聞き手としての InterActor の有効性を示した。
- (2) 聞き手としての InterActor に対してユーザが発話することで、発話の OFF 区間数の増大と全 OFF 区間の増加、それに伴い発話時間が有意に延びることを示した。
- (3) 明白に間を取る代表的な OFF 区間では InterActor の反応がある場合で OFF 区間が延びる傾向が見られた。

この発話特性の変化は、ユーザの InterActor への語りかけに対して、話の切れ目で InterActor が頷きなどの身体動作を提示することで、ユーザの音声と InterActor の身体動作が負の相関関係で相互に同調する引き込み反応が誘発されたためであり、InterActor を介したシステムが肯定的な評価を得られた要因の一つであると考えられる。

また、この InterActor が発話音声に与える効果は、人間側のユーザビリティを向上させるばかりでなく、ワードスポッティングを容易にし、音声認識率を向上させ、システム全体の信頼性を高める可能性など、コミュニケーション支援への応用が大いに期待される。

謝 辞

本研究は、科学技術振興機構 戦略的基礎研究推進事業 (CREST) の助成を受けて、著者らが提唱している「心が通う身体的コミュニケーションシステム E-COSMIC (Embodied Communication System for Mind Connection) プロジェクトの一環で行われた。

参考文献

- (1) 渡辺富夫, 大久保雅史: 身体的コミュニケーション解析のためのバーチャルコミュニケーションシステム, 情報処理学会論文誌, Vol.40, No.2, pp.670-676 (1999).
- (2) 廣瀬通孝, 小木哲朗, 玉川憲, 山田俊郎: 没入型コミュニケーションのための高臨場感ビデオパタ, ヒューマンインタフェース学会誌, Vol.2, No.2, pp.55-62 (2000).
- (3) 安田和隆, 杉田馨, 牛田啓太, 苗村健, 原島博: 透過型ビデオパタを用いたコミュニケーション・プレゼンテーション支援システム, ヒューマンインタフェースシンポジウム 2002 論文集, pp.589-592 (2002).
- (4) Shigeo Morishima: Anthropomorphic Dialog Agent Development Tool Using Facial Image Synthesis and Lip Synchronization, Human-Computer Interaction(Part2), Vol.2, pp.518-522 (2003).
- (5) Bente, G., Kramer, N. C., Blens, H.: Embodied Conversational Agents Research Questions and Applications, Universal Access in HCI, Vol.4, pp.1352-1356 (2003).
- (6) 渡辺富夫: 心が通う身体的コミュニケーションシステム E-COSMIC の開発, 機械の研究, Vol.53, No.1, pp.9-16 (2001).
- (7) 渡辺富夫, 大久保雅史, 小川浩基: 発話音声に基づく身体的インタラクションロボットシステム, 日本機械学会論文集(C 編), Vol.66, No.648, pp.251-258 (2000).
- (8) 檀原龍正, 渡辺富夫: 音声駆動型複数身体引き込みキャラクタ"InterActor"によるコミュニケーション支援, 日本機械学会 2001 年度年次大会講演資料集, Vol.III, pp.5-6 (2001).
- (9) 渡辺富夫, 大久保雅史, 中茂睦裕, 檀原龍正: InterActor を用いた発話音声に基づく身体的インタラクションシステム, ヒューマンインタフェース学会論文誌, Vol.2, No.2, pp.21-29 (2000).
- (10) Tomio Watanabe, Masashi Okubo, Mutsuhiro Nakashige and Ryusei Danbara: InterActor: Speech-Driven Embodied Interactive Actor, International Journal of Human-Computer Interaction, Vol.17, No.1, pp.43-58 (2004).
- (11) 渡辺富夫, 大久保雅史, 小川浩基, 檀原龍正: InterRobot/InterActor 音声に基づく身体引き込みインタラクションシステム, ヒューマンインタフェースシンポジウム 2001 論文集, pp.397-400 (2001).
- (12) Watanabe, T.: Voice-Responsive Eye-Blinking Feedback for Improved Human-to-Machine Speech Input, Advances in Human Factors/Ergonomics, Vol.19B, pp.1091-1096 (1993).
- (13) 檀原龍正, 渡辺富夫, 大久保雅史: 音声駆動型身体引き込みキャラクタ InterActor の多様なコミュニケーション動作の評価, ヒューマンインタフェースシンポジウム 2000 論文集, pp.101-104 (2000).
- (14) 檀原龍正, 南尾勝彦, 渡辺富夫, 大久保雅史: 音声駆動型身体引き込みキャラクタ InterActor の話者音声特性への影響, 情報処理学会第 64 回全国大会予稿集(4), pp.67-68 (2002).
- (15) Tomio Watanabe, Masashi Okubo and Hiroki Ogawa: An Embodied Interaction Robots System Based on Speech, Journal of Robotics and Mechatronics, Vol.12, No.2, pp.126-134 (2000).
- (16) 渡辺富夫: 人間-機械間の音声対話のための時間率適応化システム, 計測自動制御学会論文集, Vol.26, No.8, pp.902-907 (1990).