

交互発話における発話長と発話間隔の時間的階層性

三宅 美博*・辰巳 勇臣*・杉原 史郎*

Hierarchy between Speech Duration and Speech Interval in Alternate Speech

Yoshihiro MIYAKE*, Yushin TATSUMI* and Shiro SUGIHARA*

The purpose of this study is to clarify the temporal structure of speech dynamics in alternate speech. Experimental task is a very simple dialogue in which one speaks "A" and the other speaks "B" alternately. This alternate speech was realized between subject and machine, and subjects should focus their attention to the voice from machine-side. In this condition, two parameters, such as speech duration and speech interval, on machine-side were controlled. This revealed the mutual interaction of speech duration and the mutual interaction of speech interval in alternate speech. Temporal development of speech duration was much faster than that of speech interval in this interaction process, suggesting the temporal hierarchy between them. Furthermore, interaction of speech interval was influenced from breath process and interaction of speech duration was effected from vocalization process. These results suggest that human communication in alternate speech is a kind of hierarchical dynamic process between the control of speech duration and that of speech interval.

Key Words: alternate speech, speech duration, speech interval, temporal hierarchy, breath

1. はじめに

近年のヒューマンインタフェースには、使うことではなく、関係することを目的としたものが見受けられる。たとえばペットロボット¹⁾がそうである。ユーザは人工物との関わりとそれに基づく創出的プロセス、すなわちコミュニケーションを楽しんでいるのであり、人間のこころへの欲求が強まる昨今、このような関係性と創出性は非常に重要なパラダイムになっていくものと思われる。

われわれは、この問題を考えていく際に、人間のコミュニケーションが本来持っている、意識的な側面と非意識的な側面という二重性を無視できないと考えている。他者と関わるコミュニケーション過程を、感覚系と運動系を介するクロスフィードバックループと捉えた場合、そこに非意識的領域を介する自動的な作用関係と、意識を介する意味的な作用関係が存在することは経験的にも自然なものであろう。そして、それらが相互に影響しあうことで醸成される完全には予測できない状況が、人間に常に新しい発見を提供し、創出的コミュニケーションの楽しさを演出しているのである。

翻って従来のヒューマンインタフェース研究を見ると、そのほとんどにおいて意味づけられた情報を伝えるという意識的な明在的側面が注目され、まだ意味づけられていない情報が共有される非意識的な暗在的側面、さらにはこれら両側面の相補的な関係については十分に考慮されてこなかったよう

に思われる。われわれは、人工物とユーザのインタフェースにおいて創出的なコミュニケーションを成立させるためには、この両者の間に上記のような相補的な関係性としての二重性を積極的に導入していく必要があると考えている²⁻⁴⁾。

そして、このような観点からいくつかのヒューマンインタフェースの研究を進めており⁵⁻¹³⁾、本研究では特に交互発話を取り上げる。これは第一に、交互発話が人間のコミュニケーションの基本的形態のひとつであり、当然、同様の構造に支えられていると考えられるからである。そして第二に、従来の対話に関する研究において、そのような構造が充分には考慮されてこなかったと考えられるからである。本研究では、これらの背景に基づいて、交互発話のメカニズムを明らかにすることを目的としている。

以下、本論文の概要をまとめる。まず2章では、交互発話に関する先行研究を概観し、そこにおける問題点を整理する。つぎに3章では実験方法と実験システムを説明し、4章では実験結果を述べる。そして5章では、これらの結果を踏まえて考察を加える。特に本論文では、人間-機械系として交互発話を再構成し、その動的特性を解析する。まず、機械側が定常的に発話を生成する条件下での人間との交互発話実験を行ない、この時の人間の発話に見られる時間的な特徴を調べる。さらに、機械が生成する発話に摂動を加える実験を行ない、この時の人間の発話に見られる動的な特徴を解析する。さらに、その生理機構としての呼吸にも注目し分析を行なう。

2. 先行研究

交互発話には数多くの先行研究があるが、これらは大きく、工学的応用に主眼を置いたものと、心理計測に主眼を置いた

* 東京工業大学大学院総合理工学研究科 横浜市緑区長津田町 4259

* Tokyo Institute of Technology, 4259 Nagatsuta, Midori, Yokohama

(Received July 8, 2003)

(Revised January 6, 2004)

ものと分類できる。

まず工学的応用に主眼を置いたものには、音声案内システムのような研究がある。ここでは、質問(ユーザから機械への入力)に対する答え(機械からユーザへの出力)があらかじめデザインされシステムに実装される。このシステムは、意味づけられた情報を伝えることを目的に構成されており、コミュニケーションの意識的側面に特化した研究と言える。したがって、このような研究の枠組みでは、ユーザと人工物との間に創出的コミュニケーションを成立させることは難しい。これとは逆に、感覚系と運動系を介してユーザと人工物をリンクする中でクロスフィードバック系を構成し、そこに非意識的な作用関係を成立させようとした研究も見られる。岡田らは、身体的インタラクションに基づくアーキテクチャを用い、アドホックに発話を生成するヒューマンインタフェースを提案している¹⁶⁾。このインタフェースは、発話特性上、非意識的な作用関係を成立させていると言えるものである。

一方、心理計測に主眼を置いたものには、発話のピッチ構造と話者交代との関係を示した研究¹⁷⁾や、先行話者のピッチや発話長と後続話者のそれらとの相関を示した研究^{18), 19)}がある。これらの研究では、一方の話者のノンバーバル情報が他方話者の発話特性に影響を与えるという形式で議論がなされているため、入出力関係に則した前述の音声案内システムと類似したアプローチである。これとは別に、対面コミュニケーションにおける呼吸の引き込みに関する研究^{20), 21)}も報告されている。また発話長や間についても同様の報告があり^{22)~26)}、さらに母子インタラクションにおける話しかけと身体運動の引き込みに関する報告^{27)~29)}もなされている。二者間での呼吸と体肢運動の協調に関する知見もある³⁰⁾。これらの研究で注目されている引き込みは、一方が他方に情報を伝える構造ではなく、双方が協調して関係を創り出す構造を持っているため、コミュニケーションの非意識的な側面を適確かつ定量的に捉えたものと言える。

しかし、これらの先行研究では、コミュニケーションの意識的側面と非意識的側面に関する研究がそれぞれ独立に進められており、その両者の関係を計測し解析したものは少ない。たとえば意識的側面と深く関連する文脈情報や意味的連関が、非意識的過程に影響することを示唆する知見が報告されているくらいである^{31), 32)}。そこで、本研究では、コミュニケーションの非意識的側面と意識的側面との相補的關係を明らかにするために、両者の相互作用を解析する。具体的には、交互発話という聴覚系と発話運動系から構成されるクロスフィードバック系を取り上げる。そして、その中に意識を介する過程とそうではない非意識的な過程という二重化された作用関係が存在すると仮定し、それをコミュニケーションの二重性として捉える。これは言語等の脳高次機能を介する意味的な作用関係と、意味づけられる前の感覚運動連関レベルでの作用関係に対応するものと考えられる。そして、交互発話のメカニズムを、上記の二重性に関する仮説を踏まえ定量的に分析していくこととする。特に、発話ダイナミクスの時間的構造、および、その背景にある生理的機構としての呼吸や注意状態に注目して解析を進める。

3. 実験システム

3.1 実験における発話内容

実験において課される交互発話の内容は、一方がA(エー)他方がB(ビー)と交互に発話し、双方で協力し文字を繰り返し読み上げていくものである。これは研究の第一段階として、双方の発話内容を一音節長音に合わせることで解析を容易にするとともに、言語的な意味やコンテキストに左右されない特徴を調べることを狙っている。実験はすべて同一の発話内容で行なう。

3.2 交互発話の特徴量

実験システムを説明する前に、ここで交互発話の発話ダイナミクスに関する特徴量と呼吸に関する特徴量を、Fig.1に示すように定義しておく。

発話特徴量

<発話長>

ある話者が発話を開始してから終了するまでの時間長を「発話長(speech duration)」と定義する。たとえば、話者Aが発話を開始してから終了するまでの時間長を、「話者Aの発話長」と呼ぶこととし $Sp(A)$ と表記する。

<発話間隔>

一方の話者が発話を開始してから他方話者が発話を開始するまでの時間長を「発話間隔(speech interval)」と定義する。たとえば、話者Bが発話を開始してから話者Aが発話を開始するまでの時間長を、「話者Aによって与えられた発話間隔」と呼ぶこととし $Si(A)$ と表記する。

<間長>

一方の発話が終了してから、他方話者が発話を開始するまでの時間長を「間長(pause duration)」と定義する。たとえば、話者Bが発話を終了してから話者Aが発話を開始するまでの時間長を、「話者Aの間長」と呼ぶこととし $Pd(A)$ と表記する。

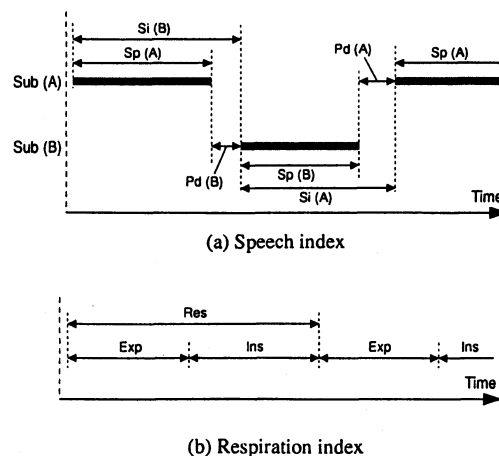


Fig.1 Feature index of alternate speech

呼吸特徴量

＜吸気時間＞

ある話者の一回の息継ぎに要する時間のうち、息を吸っている時間の長さのことを「吸気時間(inspiration duration)」と定義し Ins と表記する。

＜呼気時間＞

ある話者の一回の息継ぎに要する時間のうち、息を吐いている時間の長さのことを「呼気時間(expiration duration)」と定義し Exp と表記する。

＜呼吸周期＞

ある話者が呼気を開始してからつぎの呼気を開始するまでの時間長のことを「呼吸周期(respiration period)」と定義し Res のように表記する。

3.3 実験課題と被験者

本研究では、人間側がA(エー)機械側がB(ビー)と交互に発声する交互発話が用いられた。被験者へは、ヘッドフォンから聞こえる機械側からのB(ビー)という音声に注意を向け、機械と交互に発話を行なうように指示を与えた。このとき機械側では、20代男子学生のB(ビー)の音声を少し長めに録音しておき、それを必要な時間長だけ再生する方法で音声刺激を構成した。その上で、以下に示す2種類の実験を行なった。

実験1: 機械側の発話特徴量を固定した場合における人間側の発話と呼吸の計測と解析

実験2: 機械側の発話特徴量を変化させた場合における人間側の発話と呼吸の計測と解析

ただし被験者への視覚的影響を除くために目隠しを着用させた。また室温 18～20 度の静かな室内で実験は行なわれた。これらは被験者 5 名(20 代の健康な男子学生)の協力のもとに実施された。

3.4 実験システム

本実験で使用した計測システムの概略を Fig.2 に示す。本システムは大きく分けて、音声に関わる処理を行なうシステムと呼吸の計測を行なうシステムから構成される。

まず、音声処理システムのハードウェアとしては、ヘッドフォン(audio-technica ATC-H7COM)、ミニミキサー(FOSTEX MN06)、ノートPC(SONY PCG-803, CPU:PentiumII 233MHz, メモリ:128MB)、A/D コンバータ(ADTEK AXP-AD02, PCMCIA TypeII 準拠)を使用する。被験者に装着したヘッドフォンセットでは、そのマイクロフォンから人間の音声の入力がなされ、そのスピーカを通して機械側音声の提示がなされる。ヘッドフォンから提示する音の大きさは、被験者が音刺激を明瞭に聞き取れるレベルに調整してもらうが、この大きさは同一の被験者に対しては一連の実験を通して一定に保たれる。一方、ミニミキサーではステレオ音声による入力がなされ、人間側の音声および機械側の音声が、それぞれ異なるラインへと分離される。また入力された音声信号は、それが発声されている場合ミニミキサーに搭載されたLEDが点灯し、これをミキサー上部に取り付けられた光センサによって検出し、閾値を 0.1V とし ON-OFF の信号に変換されたデータがノートパソコンに格

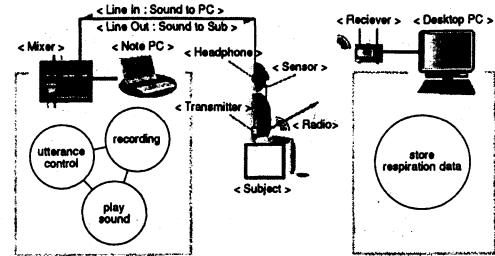


Fig.2 Experimental system

納される。

ソフトウェアは、C++Builder3(INPRISE)の環境で独自に開発したものを使用する。開発したソフトは録音部と発話制御部とから構成され、それぞれ独立に動作する。前者ではライン入力のメモリ上へのデジタル録音(サンプリング周期:20KHz, ステレオ録音)を行ない、後者ではA/Dコンバータ経由で取得された音声の有無に関する情報に基づいて発話制御を行なう。発話制御部の動作方法については、実験ごとに次節以降で詳述するが、基本的に、あらかじめ長めに録音した「B」という音声を必要な長さだけ再生する方法でさまざまな発話長や発話間隔を構成している。この再生長の調整に際しては、音声データの先頭の位置を固定し、末尾の位置を前後に制御することで実現している。また当該ソフトウェアは、精度 10ms でリアルタイム動作することができる。これらのシステムの動作フローをまとめると以下のようなになる。

＜録音部＞

1. 人間と機械の発話のミキシング
2. ミキシング音のライン入力および A/D 変換
3. 変換されたデータのメモリ上への記録

＜発話制御部＞

1. 人間と機械の発話のピーク検出および光出力
2. 光入力の電圧値変換およびデジタル化
3. 発話モデルへのデータ入力
4. 発話モデルに基づく発話制御
5. 機械の発話のヘッドフォン出力

つぎに、呼吸計測システムのハードウェアとしては、サーミスタセンサー(日本光電 TR-511G)、送信機(日本光電 XB-581Z)、受信機(日本光電 Multi Telemeter System WEB-5000)、A/Dコンバータ(Interface IBX-3119)、デスクトップPC(EPSON CPU:PentiumII 450MHz, メモリ:128MB)を使用する。サーミスタセンサーは、被験者の唇に装着され、呼気と吸気の温度差から呼吸の検出を行なう。計測されたデータは送信機を通して受信機に送信され、A/Dコンバータを経由して(サンプリング周期:100Hz)最終的にデスクトップPCに保存する。なお呼吸の検出方法としては、胸部ピックアップによる方法や鼻腔にサーミスタセンサーを装着する方法等が広く用いられている。しかし、本実験では、呼気時間および吸気時間の区別を明確にするために、被験者の唇にサーミスタセンサーを固定し、呼吸に伴う空気の流れを直接計測することとした。

3.5 データ解析方法

実験終了後、記録されたデータから、発話特徴量と呼吸特徴量を求める。発話特徴量については、録音された WAVE 形式のファイルを用いてデータ解析を行なう。データ解析用のソフトウェアは、先程と同様 C++Builder3 (INPRISE) の環境で独自に開発したものを使用した。データにはまず、10ms の精度で音圧レベルを求める処理を施す。つぎに、あらかじめ設定した音圧レベルの閾値(1dB)に基づき、有音部と無音部とに二値化する。これは音声処理システムの光センサーによる二値化と正確に対応している。最後にこれらのデータを整理し、発話長、発話間隔および間長を求める。また、呼吸特徴量については、デスクトップ PC に格納された呼吸データから呼吸に伴う温度振動のピークを求め、温度上昇の区間として呼気時間を、温度下降の区間として吸気時間を求める。呼吸停止期間は吸気時間に含まれている。さらに、呼気の開始時刻から次の呼気開始までの時間長として呼吸周期を求める。その他の解析の詳細については、その都度必要に応じて説明する。

4. 交互発話実験

4.1 実験1：機械側の発話が定常である場合

実験1では、定常的な交互発話における発話と呼吸の特徴を明らかにすることを目的とする。具体的には、機械側の発話特徴量を固定した場合における、人間側の発話特徴量と呼吸特徴量を調べる。ここでは、以下の7つの実験条件を行なった。これによって機械側の発話長 $Sp(\text{機械})$ および発話間隔 $Si(\text{機械})$ を要因とする、それぞれ1要因の被験者内計画を2種類実施する。このとき被験者は各実験条件を1回ずつ行ない、その試行の順序はランダム化されている。また、それぞれの実験時間はいずれも40秒とした。なお前節でも説明したように、 $Si(\text{機械})$ は機械によって与えられる発話間隔であり、これは人間側が発話を開始してから機械側が発話を開始するまでの時間長であることに注意する必要がある。

<実験条件>

- (1):($Sp(\text{機械})$, $Si(\text{機械})$) = (300ms, 600ms)
- (2):($Sp(\text{機械})$, $Si(\text{機械})$) = (300ms, 650ms)
- (3):($Sp(\text{機械})$, $Si(\text{機械})$) = (300ms, 700ms)
- (4):($Sp(\text{機械})$, $Si(\text{機械})$) = (300ms, 750ms)
- (5):($Sp(\text{機械})$, $Si(\text{機械})$) = (350ms, 600ms)
- (6):($Sp(\text{機械})$, $Si(\text{機械})$) = (400ms, 600ms)
- (7):($Sp(\text{機械})$, $Si(\text{機械})$) = (450ms, 600ms)

Fig.3 は、実験条件(1)において得られた発話特徴量の時間発展の一例である。図から明らかなように、機械側が設定する発話特徴量 ($Sp(\text{機械})$, $Si(\text{機械})$) が一定の条件の下では、人間側の発話特徴量 ($Sp(\text{人間})$, $Si(\text{人間})$) が機械側の特徴量に対して適応することがわかる。このとき人間側の特徴量は、それぞれ機械側の設定値である 300ms と 600ms にほぼ一致している。ただし、この一致の程度には被験者間でばらつきがあり、あくまでも平均的な性質である。その他の実験条件(2)~(7)においても、それぞれ同様の傾向が観察された。この結果より、本実験

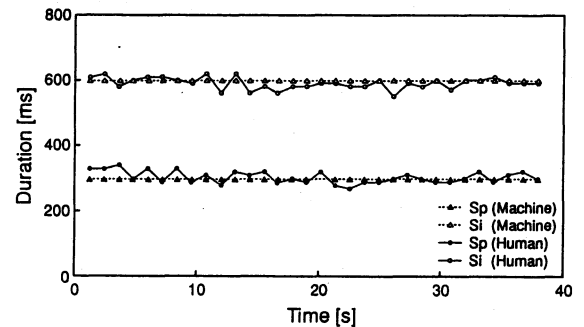
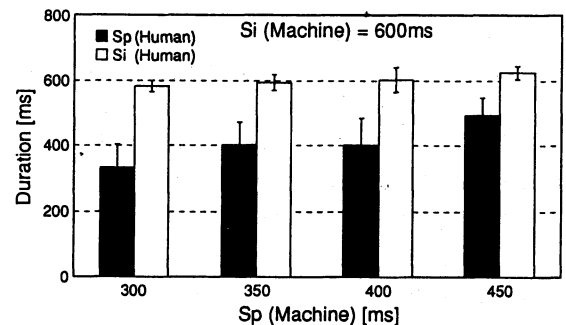
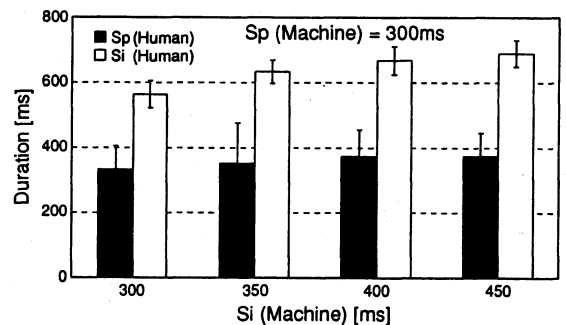


Fig.3 Temporal development of speech index



(a) Effect of $Sp(\text{Machine})$



(b) Effect of $Si(\text{Machine})$

Fig.4 Relationship between speech indexes

条件において、人間の発話長および発話間隔は、機械側が設定する発話特徴量に対してその時間長を一致させる傾向があり、それぞれ引き込まれることが示唆された。これらの性質は以下の Fig.4a,b から確認される。

Fig.4a は、機械によって与えられる発話間隔 ($Si(\text{機械})$) を 600ms で固定した条件下での、機械の発話長 ($Sp(\text{機械})$) と人間側の発話特徴量 ($Sp(\text{人間})$, $Si(\text{人間})$) の関係を示したものである。ここでは、各試行において開始 10 秒以降に観測された人間側の発話特徴量の被験者 5 名に対する平均値を示している。ただし、実験システムの精度が 10ms であることから、平均値についても同じ精度で求めることとする。この図から明らかなように、機械の発話長 ($Sp(\text{機械})$) は人間の発話長 ($Sp(\text{人間})$) に影響するが、人間によって与えられる発話間隔 ($Si(\text{人間})$) にはあまり影響しないことがわかる。ここで $Sp(\text{機械})$ を要因とする 4 水準の 1 元配置のデータに対して繰り返しの無い 1 元配置の分散分析を行なった結果、 $Sp(\text{機械})$ から $Sp(\text{人間})$ が有意に影響を受けることが確認された ($F(3,16) = 4.54$, $p < 0.02$)。また、同

様の手法によって、Sp(機械)からSi(人間)へは顕著な影響が認められない($F(3,16) = 2.39$, $p > 0.1$)ことも示された。これらの結果より、定常的な交互発話においては、発話長間の作用関係が強い一方、発話長から発話間隔への影響は弱いことが示された。

ただし、本結果は1要因被験者内計画を2種類独立に実施して得たものであり、2要因被験者内計画によって得られたものではないことに注意しておく必要がある。この制約は本節のすべての結果において考慮されなければならない。

Fig.4bは、機械の発話長(Sp(機械))を300msで固定した条件下での、機械によって与えられる発話間隔(Si(機械))と人間側の発話特徴量(Sp(人間), Si(人間))の関係を示したものである。この場合も、各試行において開始10秒以降に観測された人間側の発話特徴量の平均値が用いられている。この図から明らかなように、機械によって与えられる発話間隔(Si(機械))は人間の発話長(Sp(人間))にあまり影響しないが、人間によって与えられる発話間隔(Si(人間))には影響することがわかる。ここでSi(機械)を要因とする1元配置のデータに繰り返しの無い1元配置の分散分析を行なった結果、Si(機械)に対してSp(人間)は有意な影響を受けないことが示された($F(3,16) = 0.25$, $p > 0.8$)。また、同様の手法によって、Si(機械)に依存してSi(人間)は有意な差を生じることが確認された($F(3,16) = 9.16$, $p < 0.01$)。これらの結果より、定常的な交互発話においては、発話間隔間の作用関係が強い一方、発話間隔から発話長への影響は弱いことが示された。

Fig.5aは、Fig.3のデータに対応する呼吸波形と人間の発話を示している。縦軸はサーミスタセンサーで検出された呼吸に伴う空気流の温度変化を示しており、傾きが正の場合が吸気状態(ただし呼吸停止状態を部分的に含む)、負の場合が呼気状態であることに対応している。この結果のように発話回数と呼吸回数が一致し、1回の発話で1回の呼吸を行なう関係が全被験者において観察された。Fig.5bは、このときの呼吸特徴量の時間発展である。図から明らかなように、機械側の発話特徴量(Sp(機械), Si(機械))が一定であれば、人間側の呼吸特徴量(Res, Ins, Exp)もほぼ時間的に一定であることがわかる。その他の実験条件(2)~(7)においても、それぞれ同様の傾向が観察された。

Fig.6aは、機械によって与えられる発話間隔(Si(機械))を600msで固定した条件下での、機械の発話長(Sp(機械))と呼吸特徴量(Res, Ins, Exp)の関係を示したものである。この場合も、各試行において開始10秒以降に観測された呼吸特徴量の被験者5名に対する平均値を示している。この図から明らかなように、機械の発話長(Sp(機械))は、呼吸特徴量のいずれにも影響しないことがわかる。Fig.4aと同様の統計的手法によると、Sp(機械)に対するResに有意な差は認められなかった($F(3,16) = 2.01$, $p > 0.15$)。Insについては $F(3,16) = 0.21$, $p > 0.8$ であり、Expについても $F(3,16) = 0.90$, $p > 0.4$ であり、いずれもSp(機械)からの有意な影響を受けないことが確認された。これらの結果より、発話長間の作用関係においては、呼吸との

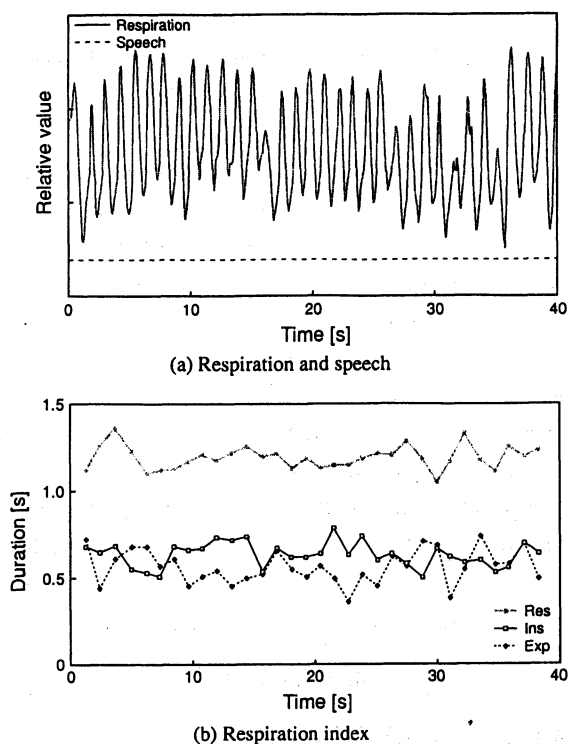


Fig.5 Temporal development of respiration index

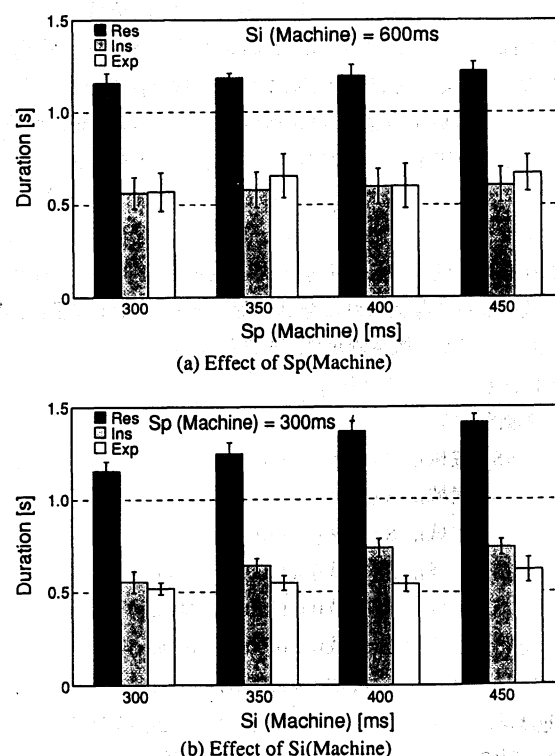
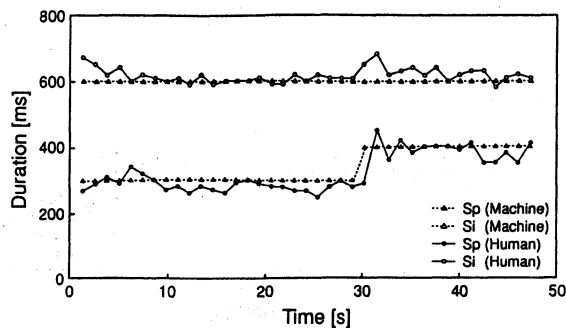


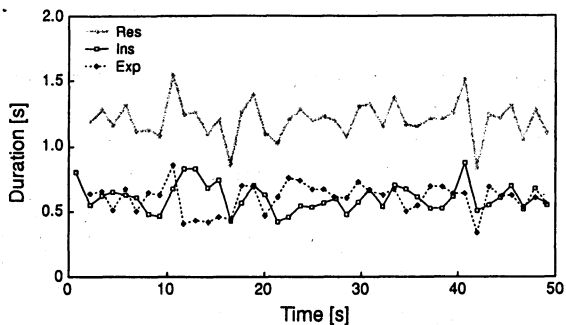
Fig.6 Relationship between speech index and respiration index

関係が弱いことが示された。

Fig.6bは、機械の発話長(Sp(機械))を300msで固定した条件下での、機械によって与えられる発話間隔(Si(機械))と呼吸特徴量の関係を示したものである。この図から明らかなように、機械によって与えられる発話間隔は、呼吸周期および吸気時間に影響するが、呼気時間には影響しないことがわかる。5名の

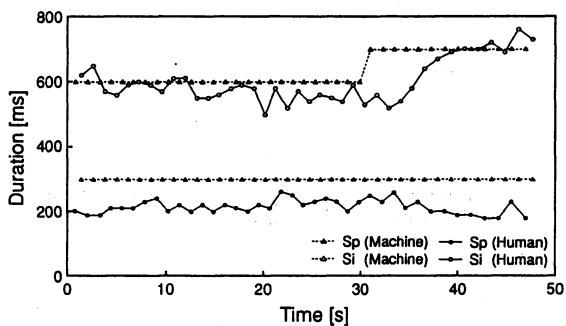


(a) Speech index

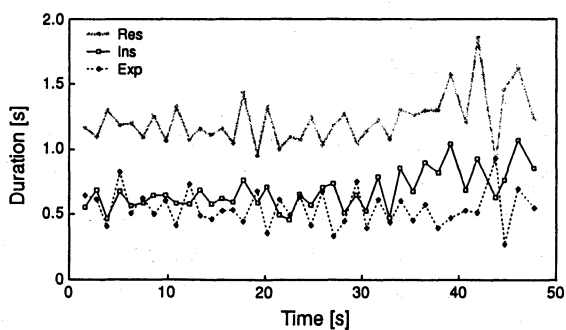


(b) Respiration index

Fig.7 Temporal development of speech index and respiration index



(a) Speech index



(b) Respiration index

Fig.8 Temporal development of speech index and respiration index

被験者に対して Fig.4b と同様に統計的に調べてみると, Si(機械)に依存して Res に有意な影響が現れることが確認された($F(3,16) = 24.22$, $p < 0.01$). また, そのうちの4名の被験者に対しては, Ins への影響において有意差が認められ($F(3,12) = 20.25$, $p < 0.01$), Exp に対しては顕著な影響が認められなかった($F(3,12) = 2.90$, $p > 0.1$). これらの結果より, 発話間隔の作

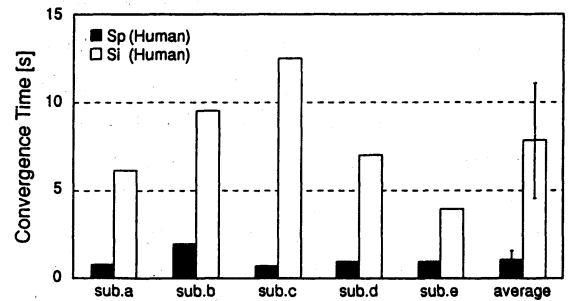


Fig.9 Convergence time of speech index

用関係では, 呼吸周期, 特に吸気時間との関係が強いことが示された. なお, 残りの1名の被験者は呼気時間を増加させることで呼吸周期の調整を行っていることが確認された. この被験者では, 2回の発話で1回の呼吸を行なっている箇所も一部見られた.

4.2 実験2: 機械側の発話に摂動が含まれる場合

実験2では, 交互発話の遷移過程における発話と呼吸の特徴を明らかにすることを目的とする. 具体的には, 機械側の発話特徴量を変化させた場合における, 人間側の発話特徴量と呼吸特徴量の対応する変化を調べる. ここでは, 下記の実験条件(8),(9)を50秒間行なうものとし, 摂動は実験開始30秒後に加えることとした. ただし, 被験者および使用した実験システム, 実験環境, その他の条件は実験1と同様である.

<実験条件>

(8):(Sp(機械), Si(機械))

= (300ms, 600ms) → (400ms, 600ms)

(9):(Sp(機械), Si(機械))

= (300ms, 600ms) → (300ms, 700ms)

Fig.7a は, 実験条件(8)において得られた発話特徴量の時間発展の一例である. 図から明らかなように, 機械によって与えられる発話間隔(Si(機械))を固定した条件下で, 機械の発話長(Sp(機械))を増加させたとき, 人間の発話長(Sp(人間))が顕著に増加することがわかる. しかも, このとき人間の発話長は機械の発話長に向かって急速に収束する傾向が見られた. Fig.7b は, 同条件において得られた呼吸特徴量の時間発展であり, いずれにおいても対応した変化が認められなかった.

Fig.8a は, 実験条件(9)において得られた発話特徴量の時間発展の一例である. ここでは機械の発話長(Sp(機械))を固定した条件下で, 機械によって与えられる発話間隔(Si(機械))を増加させたとき, 人間によって与えられる発話間隔(Si(人間))が顕著に増加することがわかる. このとき人間の発話間隔は, 機械の発話間隔に向かってゆっくりと収束する傾向も見られた. Fig.8b は同条件において得られた呼吸特徴量の時間発展である. 図から明らかなように, 人間側では呼吸周期(Res)と吸気時間(Ins)がゆっくりと増加する傾向が認められた.

これらの結果は, 実験1の定常状態で観察された発話長および発話間隔の作用関係が交互発話の遷移過程においても成立していることを示している. さらに, Fig.7aの遷移過程における発話長とFig.8aの遷移過程における発話間隔は, その変化速度

において異なる時間スケールを持つことが示唆される。そこで、実験条件(8)における発話長の収束時間と実験条件(9)における発話間隔の収束時間を、5名の被験者に対する統計的手法によって比較した。なお、収束時間は摂動付加後の発話長の変化開始時刻から定常値に到達するまでの時間としている。ここでは、摂動付加前10秒間(時刻20~30秒)の平均値を上回る点を変化開始時刻とし、計測終了前の10秒間(時刻40~50秒)の平均値を上回る点を定常値に収束した時刻として定義する。Fig.9に示すように、収束時間は被験者間でばらつきが大きい。平均値としてはSp(人間)とSi(人間)の収束時間は、それぞれ1.11秒と7.86秒であり、しかも両者の間に有意差が認められた($t(4) = 4.64$, $p < 0.01$)。

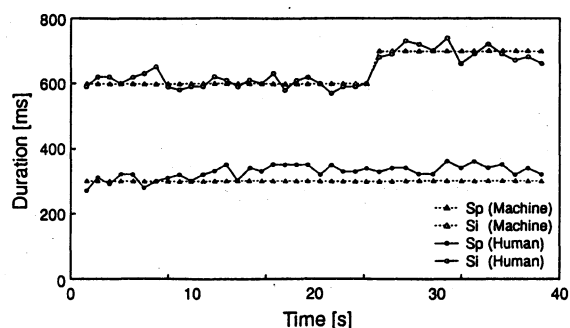
5. 考察

実験1の結果より、交互発話における発話間隔の引き込みおよび発話長の引き込みの傾向が確認された。このことに基づいて、発話間隔から発話間隔への作用および発話長から発話長への作用の強いことが明らかにされた。それに比して、発話間隔から発話長への作用および発話長から発話間隔への作用は弱いことも示された。これらの結果は、発話間隔に関する作用関係と発話長に関する作用関係が、交互発話において異なる種類の作用関係として存在していることを意味している。

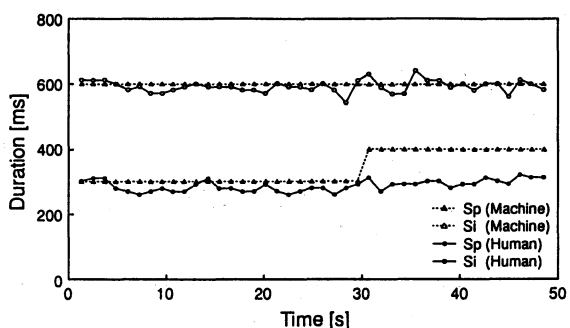
また、発話間隔に対しては、呼吸周期と吸気時間の相関が非常に高く、呼吸時間との相関は低かった。発話長については、呼吸特徴量のいずれとも相関は低かった。発声するために声帯(声門)を振動させることと、呼吸そのものを行なうことは別の制御機構に属することを考慮すれば、これらの結果は、発話間隔は呼吸機構からの影響を強く受け、発話長は発声機構からの影響を強く受けていることを示唆するものである。ただし、本実験の呼吸計測方法の範囲内では、吸気状態と呼吸停止状態を区別できない点は考慮しておく必要がある。

実験2では交互発話における遷移過程の性質が調べられた。そして、このような条件下においても、発話間隔の作用関係および発話長の作用関係がそれぞれ強いことが示された。また、呼吸特徴量との関係も実験1と同様であり、呼吸周期と吸気時間は発話間隔と相関を持って変化していた。さらに、そのような遷移過程の変化速度の時間スケールについても、発話長と発話間隔の間での違いが明らかにされた。つまり、発話長は短い時間スケールで応答し、発話間隔はそれに比してゆっくりと応答するのである。

このような発話間隔と発話長の時間発展における時間スケールの違いを説明できる可能性として、「注意」について考えてみる事ができる。ある刺激に対して被験者の応答が相対的に短い時間で観察されたとすれば、その対象(刺激)に注意が向けられていることが示唆されるからである。このような傾向は視覚情報の処理過程の加速として観察されており³³⁾、さらに、本実験と類似した感覚運動連関である同期タッピング課題の位相補正でも報告されている³⁴⁾。本実験では、被験者に機械側のB(ビー)という発話音声に注意を向けるように指示してお



(a) Perturbation of speech interval



(b) Perturbation of speech duration

Fig.10 Temporal development of speech index

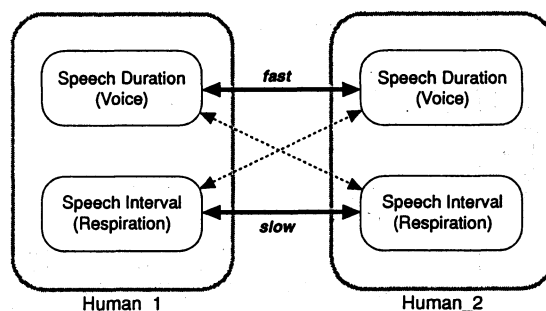


Fig.11 Temporal hierarchy of alternate speech

り、このような条件によって発話長に関わる処理が発話間隔に関わる処理に比して加速され、上記のような時間的階層性を生じたものと推測される。

このことは、被験者の注意を発話ではなく発話間隔に向けられる条件下で、同様に変化速度を計測することによって確認される。結果の一例を Fig.10a に示すが、このときは機械側の発話間隔を変化させると被験者側の発話間隔も急速に追従するのである。これを Fig.8a と比較すれば、両者の発話間隔の変化速度に時間スケールの違いが存在することがわかるであろう。このことは、本実験において被験者の注意状態が適切にコントロールされていることを示すだけではなく、実験2における発話長と発話間隔の変化速度における時間スケールの違いが、被験者の注意に依存していることを意味する。

さらに、Fig.10b のように、発話間隔に注意を向けたままで、機械側の発話長を変化させると被験者側の発話長は変化しないことも示された。このことは、発話長変化には注意が不可欠であることを意味しており、図 8a のように、注意が発話間隔に

向いていなくても発話間隔がゆっくりと変化することと対照的である。このことから発話長変化は発話間隔に比して注意からの影響を相対的に強く受けていることが予想される。

以上の結果より、発話長の作用関係と発話間隔の作用関係における時間スケールの違いは、主として被験者の注意状態に起因することが示唆された。注意は被験者の意識状態と密接な関係を有しているので³⁵⁾、このことは発話長の作用関係がコミュニケーションの意識的側面に対応し、発話間隔の作用関係が非意識的なコミュニケーションに対応する可能性を示唆するものである。このことから交互発話においてもコミュニケーションの二重性の存在が推測される。ただし本研究においては、被験者の意識状態そのものを計測した訳ではないので、より厳密には被験者の内観との関連付けが必要であろう³⁶⁾。

すでに、呼吸がコミュニケーションにおける非意識的な側面と関係していることは示唆されてきたが^{20), 21), 37)}、本実験においても発話間隔と呼吸の相関が高かったことは上記の考察を支持する結果といえる。また、発声に関連して、吐き出す空気流に対してどこからどこまで声帯を振動させるかという発話制御は意識的なものと考えられ、これも同様に上記の考察を支持するものである。

したがって、人間2人の交互発話における発話特徴量の間には、Fig.11のような二重化された作用関係が存在するものと予想される。発話長は発声機構からの制御を受けており、しかも、意識が関与する注意の影響下におかれている可能性がある。一方、発話間隔は注意からの影響の弱い非意識的領域に属する可能性があり、呼吸過程という身体的インタラクションからの影響を強く受けている。そして、これらのダイナミクスは時間スケールにおける階層性を構成していると考えられる。

6. おわりに

本研究においては、交互発話というコミュニケーションに注目し、それが二重化された階層的な作用関係から構成されていることを明らかにした。一方は意識的な側面であり、他方は非意識的な側面である。ただし、これらの2種類の作用関係の関係については、今回の実験の範疇では十分に解明することができなかった。これらの二重化された作用関係は独立に存在するのではなく、時間的階層性を介して動的連関を持ちつつ相補的に機能するものと予想されるが、その問題は今後の課題とせざるをえないだろう。しかし、このような二重性として人間のコミュニケーション機構を解明できる可能性が示されたことは重要である。

なお、本研究で用いた交互発話系は日常の対話コミュニケーションと比較すると特殊な状況にある。しかし、ここで得られた結果は、人間-機械系において機械が人間の対話コミュニケーションへ介入できる可能性を示しており、コミュニケーション系の再構成へと道を拓くものである。また本研究の応用領域としては、音声対話インタフェースの発話タイミングの制御が考えられる。これは、発話タイミングの生成が、対話の中で時々刻々と創り出されてゆくもので、従来のルールベースに

よる記述に馴染まないと考えられるからである。たとえば、対話型のナビゲーターシステムや、電話を介した応対システムといった領域が考えられる。

今後これらを実際に人工物として構築するためには、対話の意味やコンテキストなども考慮する必要があり、まだまだ残された問題点も多い。ただ、本研究において明らかにされた発話間隔と発話長の階層的な作用関係は、創出的コミュニケーションに関わる非常に興味深い問題である。このような二重性に基づく対話コミュニケーションの研究の進展が、これから一層求められるであろう。

参考文献

- 1) AIBO: <http://www.aibo.com/> (1999)
- 2) 清水博: 生命知としての場の論理, 中公新書 (1996)
- 3) 清水博: 新版 生命と場所, NTT 出版 (1999)
- 4) 清水博: 共創と場所, 場と共創, NTT 出版, 23/178 (2000)
- 5) マイケル・ボランニー: 暗黙知の次元, 伊藤敬三訳, 紀伊国屋書店 (1980)
- 6) 三宅美博: 「生命」における設計, 現代思想, 25-6, 301/317 (1997)
- 7) 三宅美博: 「二中心モデル」とインタフェース表現, 日本ファジィ学会誌, 9-5, 637/647 (1997)
- 8) 三宅美博: コミュニカビリティと共生成, 場と共創, NTT 出版, 339/397 (2000)
- 9) Y. Miyake and J. Tanaka: Mutual-entrainment-based internal control in adaptive process of human-robot cooperative walk, Proc. 1997 IEEE Int. Conf. on Systems, Man and Cybernetics (SMC'97), Orlando, U.S.A. 293/298 (1997)
- 10) Y. Miyake, T. Miyagawa, and Y. Tamura: Internal observation and mutual adaptation in human-robot cooperation, Proc. 1998 IEEE Int. Conf. on Systems, Man and Cybernetics (SMC'98), San Diego, U.S.A. 3685/3690 (1998)
- 11) Y. Miyake and T. Miyagawa: Internal observation and co-generative interface, Proc. of 1999 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC'99), Tokyo, Japan, I-229/I-237 (1999)
- 12) Y. Tatsumi and Y. Miyake: Hierarchy and Fluctuation in Time Structure of Spoken Dialogue, Proc. of 1999 IEEE Int. Conf. on Systems, Man and Cybernetics (SMC'99), V-1040/V-1044, Tokyo (1999)
- 13) 三宅, 宮川, 田村: 共創出コミュニケーションとしての人間-機械系, 計測自動制御学会論文集, 37-11, 1087/1096 (2001)
- 14) 武藤, 三宅: 歩行介助を目的とする人間-ロボット協調系における共創出過程の解析, 計測自動制御学会論文集, 38-3, 316/323 (2002)
- 15) 高梨, 三宅: 共創型介助ロボット "Walk-Mate" の歩行障害への適用, 計測自動制御学会論文集, 39-1, 74/81 (2003)
- 16) 岡田美智男: 社会的な相互行為とそのリアリティを支えるもの, 身体性とコンピュータ, bit 別冊, 岡田美智男, 三嶋博之, 佐々木正人(編), 共立出版, 220/232 (2000)
- 17) 小磯, 堀内, 土屋, 市川: 言語的・韻律的情報を利用した発話の終了/継続の予測, 人工知能学会全国大会第10回論文集, 407/410 (1996)
- 18) 垣田邦子: 話し言葉における話者間の発話リズムの相互作用に関する研究—文長とポーズ長に見られる話者間相互作用について—, 文部省科研費重点研究平成5年度成果報告書, 67/73 (1994)
- 19) 垣田邦子: 簡単な「質問—答」形式の対話における F0 の話者間相互作用, 音響学会秋季講演論文集, 2-P-2, 305/306 (1995)
- 20) 渡辺富夫, 大久保雅史: コミュニケーションにおける引き込み現象の生理的側面からの分析, 情報処理学会論文誌, 39-5, 1225/1231 (1998)
- 21) T. Watanabe and M. Okubo: Evaluation of the Entrainment Between a Speaker's Burst-Pause of Speech and Respiration and a Listener's

Respiration in Face-to-Face Communication, Proc. of 6th IEEE Int. Workshop on Robot and Human Communication (ROMAN'97), 392/397 (1997)

- 22) 綿貫, 外川: 対話における感情の変化の解析, 音声言語処理, 7-13, 79/84, 1995.
- 23) 中村敏枝: '間'の感性, 感性情報処理, 電子情報通信学会(編), オーム社, 151/162 (1994)
- 24) 長岡, 小森, 中村: 自由対話における話者交代の潜時と呼吸の関連, ヒューマンインタフェースシンポジウム 2000 論文集, 311/314 (2000)
- 25) 中村敏枝: ノンバーバルコミュニケーション - 「間」が伝達する感性情報 -, ヒューマンインタフェースシンポジウム 2001 論文集, 87/88 (2001)
- 26) 杉藤美代子: 日本人の声, 和泉書院 (1994)
- 27) W.S.Condon and L.W.Sander: Neonate Movement is Synchronized with Adult Speech, Science, 183, 99/101 (1974)
- 28) C.Trevarthen: Musicality and the intrinsic motive pulse: evidence from human psychology and infant communication, Musica Scientiae, Special Issue 1999-2000, 155/211 (2000)
- 29) S.N.Malloch: Mothers and infants and communicative musicality, Musica Scientiae, Special Issue 1999-2000, 29/54 (2000)
- 30) 高瀬, 古山, 三嶋, 春木: 二者間の呼吸と体肢運動の協調, 心理学研究, 74-1, 36/44 (2003)
- 31) E.Saltzman and J.A.S.Kelso: Skilled actions: A task-dynamic approach, Psychological Review, 94, 84/106 (1987)
- 32) Y.Furuyama: Gestural interaction between the instructor and the learner in origami instruction, In D.NcNeill (Ed.), Language and Gesture, Cambridge University press, Cambridge, 99/117 (2000)
- 33) 宮内, 沖田: 注意システムと意識, 脳と意識, 苅坂直行(編), 朝倉書店, 146/187 (1997)
- 34) M.H.Thaut, R.A.Miller and L.M.Schauer: Multiple synchronization strategies in rhythmic sensorimotor tasks: Phase vs period correction, Biological Cybernetics, 79, 241/250 (1998)
- 35) 苅坂直行: ワーキングメモリと意識, 脳とワーキングメモリ, 苅坂直行(編), 京都大学学術出版会, 1/20 (2000)
- 36) B.H.Repp: Phase correction, phase resetting, and phase shifts after subliminal timing perturbation in sensorimotor synchronization, Journal of Experimental Psychology: Human Perception and Performance, 27, 600/621 (2001)
- 37) 古浦一郎: 心理学的考察「いきが合う」, 北大路書房 (1990)

「著者紹介」

三宅 美博 (正会員)



1989年東京大学大学院薬学系研究科博士課程修了(薬学博士). 金沢工業大学情報工学科助手, 講師, 助教授を経て, 96年より東京工業大学大学院総合理工学研究科知能システム科学専攻助教授. 99年よりミュンヘン大学客員教授併任. 主として生命的自律性の研究に従事, 生物物理学会, IEEEなどの会員.

辰 巳 勇 臣



1992年東京大学工学部機械工学科卒. 2000年東京工業大学大学院総合理工学研究科知能システム科学専攻博士前期課程修了. 現在, 日本電気(株)インターネットシステム研究所勤務.

杉 原 史 郎



1999年同志社大学工学部知識工学科卒. 2002年東京工業大学大学院総合理工学研究科知能システム科学専攻博士前期課程修了. 現在, オムロン(株)勤務.