

特集 「会話情報学」

会話の分析とモデル化

Analysis and Modeling of Conversation

植田 一博
Kazuhiro Ueda
東京大学大学院総合文化研究科
Department of System Sciences, The University of Tokyo.
ueda@gregorio.c.u-tokyo.ac.jp

吉川 左紀子
Sakiko Yoshikawa
京都大学大学院教育学研究科
Graduate School of Education, Kyoto University.
150027@sakura.kudpc.kyoto-u.ac.jp

伝 康晴
Yasuharu Den
千葉大学文学部
Faculty of Letters, Chiba University.
den@cogsci.L.chiba-u.ac.jp

長岡 千賀
Chika Nagaoka
京都大学大学院教育学研究科（日本学術振興会特別研究員）
Graduate School of Education, Kyoto University.
nagaoka@educ.kyoto-u.ac.jp

大本 義正
Yoshimasa Ohmoto
東京大学大学院総合文化研究科
Department of System Sciences, The University of Tokyo.
ohmoto9@dolphin.c.u-tokyo.ac.jp

榎本 美香
Mika Enomoto
千葉大学大学院自然科学研究科
Graduate School of Science and Technology, Chiba University.
menomoto@cogsci.L.chiba-u.ac.jp

Keywords: automatic discrimination of lies, dialogue, embodied synchrony, eye-gaze, multi-modal conversation, multi-party conversation, non-next-speaking participants, nonverbal behavior, reading of intentions.

1. はじめに

会話情報学では、会話現象をコンテンツとインタラクションの両面から捉えることを目指しているが、東京大学、京都大学、千葉大学の三つのグループは、それぞれ独自に、認知心理学・心理言語学のアプローチに基づいて、特にインタラクションの側面から会話現象を捉える研究を行っている。

[Grice 69]以降、コミュニケーション研究は意図的な情報伝達を主な研究対象としてきたといえよう。しかしながら、日常的なコミュニケーション場面では、話し手が伝えようとは意図していない情報が伝わるのがしばしば起こる。さらに、受け手によって誤って解釈された話題に沿って、会話が展開されることもある。会話がこのような性質をもつ以上、会話現象をインタラクションの立場から分析するには、会話に参加しているあるメンバによって表出された発話や行動に、ほかのメンバがいかに注目しアクセスするかを重視する立場、すなわち（意図的な情報伝達よりも）表出された行動の連鎖として会話を捉える立場があり得る。しかも、他者行動へのアクセスは、意識的というよりは無意識的・自動的になされている可能

性がある。このような自動的な処理に大きな役割を果たすのは、テキストとして書き表される言語情報よりもむしろ、視線、表情などの非言語情報（韻律などのパラ言語情報も含む）だと考えられる。

以下の三つの章で紹介する研究は、いずれも、人間同士のコミュニケーションをそれに関わる者によって表出された行動の連鎖として捉え、かつその中で非言語情報が会話を成立させる社会的シグナルとしていかなる意味もっているのかを明らかにしつつある研究である。2章ではカウンセリングにおける巧拙を非言語行動の同調性の生起という観点から分析した京大グループの研究を、3章ではマルチモーダル三人会話の収録とそこでの円滑な話者交替を視線行動の共起という観点から分析した千葉大グループの研究を、4章では会話において隠べいしようとしても漏えいしてしまう情報（意図）を人はどのような非言語情報から判断しているのかを分析した東大グループの研究を紹介する。

2. 対話における聞き手の役割：聞き手の非言語行動

対話は、音声（言語）チャンネルおよび表情・しぐさを

中心とする非言語チャネルによる、二者間のコミュニケーションであり、人が日常場面で最も頻繁に利用するコミュニケーション形態の一つである。対話は、各自のもつ情報を相手に伝えることや、話題に対する感情や評価を相互に伝え合い共有することに加えて、新しい問題解決法を思いつく、新たな自己認識・他者認識に至るといった創造的・発見的な機能も有しており、知・情・意すべての心的活動が関わる高次で複雑な相互行為であるといえる。

こうした対話の多様な側面が十全に機能するうえで、話し手の発話動機の強さや言語・非言語表現のスキルに加えて、適切なタイミングで発信される聞き手からの応答が重要な要素であると考えられる。我々の研究課題は、話し手の発話が円滑に生成され、話し手にとって「話しやすい」と感じられる対話のもつ特性について、話し手・聞き手間の身体同調性や、聞き手の言語・非言語的応答特性に着目して心理実験により検証することである。題材としたのは、心理臨床の面接（カウンセリング）における、クライアントとカウンセラの対話である。

カウンセリングは、クライアントの心理的な困難や障害の克服を目的として、トレーニングを積んだ聞き手であるカウンセラとの間で行われる対話である。カウンセリングは、発話を促す主要な動機が、話し手（クライアント）の心理的問題（「悩み」）の解決であり、感情的・内省的な要素が発話内容に多く含まれること、長期間にわたって、長時間の対話（50分間/1面接）が続けられること、話し手と聞き手の役割が固定しており、カウンセラは「聞き手の専門家」という立場で対話に参加すること、という特徴があり、日常対話とは異なる性質をもっている。しかし、カウンセリングにおける優れた聞き手（カウンセラ）の応答は、話し手の発話を促進し、話し手の思考を「混乱した状態」から「整理された状態」へと修正する援助機能をもつことなどが経験的に知られており、聞き手としてのカウンセラの応答を、非言語行動も含め詳しく分析することによって、日常対話を促進する聞き手の応答様式の特長性についても重要な手掛かりが得られることが期待される。

2・1 同調性への着目

本研究では、非言語行動が対話参加者間で同期したり、音声パターンが類似したりするという同調性の現象に焦点を当てて、対話の円滑さにおける聞き手の役割について検討した。同調性は、一方の話者の発話のオンオフと聞き手の頭部や腕などの運動が同じ時点で起こる、あるいは、姿勢や表情、声の大きさ、高さ、ポーズや反応潜時（相手が話し終わってから自らが話し始めるまでの時間長）のとり方が類似するなど、さまざまなコミュニケーションチャネルに多様な様式で表れる[大坊 98]。

対話場面における非言語行動の同調性は、常に観察されるわけではなく、対話者の共感性や受容性などの要因

により変化する。例えば、意見が対立する二者が、相手の意見を受け入れながら妥協策を見いだそうと対話する場合には反応潜時が同調性を示すのに対して、相手の意見を受け入れず自分の意見を主張しようと対話する場合には同調性を示さない[長岡 03]。また、二人で旅行プランを練るといった協調的な対話では、良好な対人関係の話者同士ほど、互いの身体動作の同期が観察される[Bernieri 96]。このことから、カウンセリング場面における同調性は、クライアントとカウンセラ（聞き手）間の信頼性などの心理要因と相関が高いことが予測できる。

2・2 スクールカウンセリング映像を用いた心理評定

我々は、スクールカウンセリング映像を用いた心理評定実験により、生徒（クライアント）と教師やスクールカウンセラ（聞き手）の非言語行動の同調性が、カウンセリングの質や相談相手の信頼感と関連するかどうかを検証した。刺激として、教師教育教材『学校教育とカウンセリング』（宮本・山田監修）から事例 1（女子中学生と女性教師）と事例 2（男子中学生と女性スクールカウンセラ）を監修者の使用許可を得たうえで用いた。いずれも模擬カウンセリングを撮影したもので、専門家のコメント（DVD に収録）によれば、事例 1（収録時間 15 分）は「生徒が十分に内観を表現できていない」のに対し、事例 2（収録時間 27 分）は「生徒とカウンセラとのラポール（良好な人間関係）がすでにある程度存在している」、「全体に温かい雰囲気」の「良いカウンセリングの例」である。今回の心理評定に用いた映像は、各事例における冒頭（着席後約 30 秒）、相談内容の聞終り（30 秒）、その直後（30 秒間）、最後（最後の 1 分間を除く約 30 秒）の 4 場面であり、8 映像（4 場面×2 事例）であった（図 1）。

まず音声をオフにして映像を提示し、二者間の身体動作の同調性を 24 名の実験参加者に 10 段階で評定させた[Bernieri 88]。また別な機会に、9 名の実験参加者に対して、映像を見せずに音声のみ提示し音声特性の同調性（声の力強さの類似度、発話速度の類似度、合いの手・掛合いの調和の 3 項目、各 10 段階）、およびカウンセリングの質に関する心理評価を求めた。心理評価では、対話の「円滑さ」の程度、および生徒が抱く聞き手に対する信頼感の強さの 2 項目（各 10 段階）を評定させた。

結果から、身体動作の同調性と心理評価の間（円滑さ、信頼感の順に、 $r = 0.97$, $r = 0.89$, とともに $p < 0.01$ ）、および声の力強さの類似度と心理評価の間（順に、 $r =$

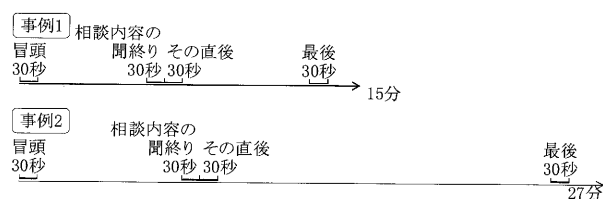


図 1 刺激の切出し箇所

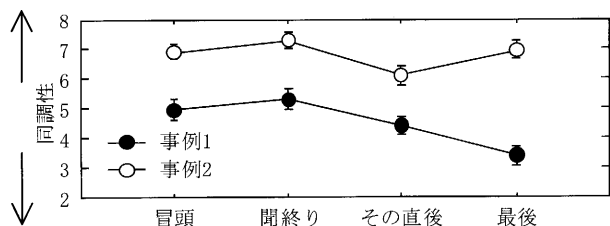


図2 時間経過による身体動作の同調性の変化。
エラーバーは標準誤差

0.88, $p < 0.01$; $r = 0.75$, $p < 0.05$), 合いの手・掛合いの調和の評定値と心理評価との間に有意な相関があることが示された ($r = 0.93$, $r = 0.87$, ともに $p < 0.01$).

また、身体動作の同調性と声の力強さの類似度、合いの手・掛合いの調和のそれぞれの間に、有意な高い相関が認められた。この結果は、身体動作に見られる同調性と声の大きさや掛合いの同調性が共起関係であることを示す。

図2は、カウンセリングの時間経過と身体動作の同調性の変動との関係を示したものである。「生徒とカウンセラとのラポールがすでにある程度存在している」、「良いカウンセリングの例」とコメントされる事例2では、一貫して比較的高い同調性が見られるのに対して、「生徒が十分に内観を表現できていない」とコメントされる事例1では、相対的に同調性が低く、特にカウンセリングの最後において顕著に同調性が低下することが示されている。カウンセリングの時間経過とともにクライアントとカウンセラの間で姿勢の同調性が高くなることを示す先行研究 [Charny 66] と本結果を考え合わせると、非言語行動の同調性は、時間経過に伴って刻々と変化するカウンセリング対話の質や、クライアントとカウンセラの心的距離などの特性を敏感に反映していることが推測される。今回の検討によって、非言語行動、音声に見られる二者間の同調性がカウンセリングの質やクライアントとカウンセラの関係性の質を反映する指標として利用できる可能性が示唆された。

今回は、同調性について、身体の動きや音声の大きさやピッチなどの物理的特性をそのまま計測・分析するのではなく、心理評価という方法を用いた。今回の結果を手掛かりに、同調性を知覚させる物理的要因を特定し、それぞれの指標の関係を明らかにすることが今後の課題である。例えば、クライアントとカウンセラの頭部の振れ、身振り、表情、視線、視線交錯、応答の韻律、発話速度、反応潜時などを分析対象とすることが考えられる。

臨床心理学では、カウンセラがクライアントとの間に適度な対人距離を維持し、共感的に応答することによって、話し手の「深い」発話が引き出されるといわれている。応答の巧拙によっては、対話の持続が困難になるという意味で、聞き手としてのカウンセラの役割は重要である。カウンセリングにおける言語応答の特徴を、非言語行動やその同調性と合わせて時系列的に分析することによ

て、二者間の対話を促進する聞き手の応答の特性が明らかになることが期待される。

3. 多人数マルチモーダル会話の収録と分析

3.1 目的

千葉大グループでは、日常場面における会話インタラクションの仕組みを明らかにすることを目的とし、人間同士の自然な会話を収録し、主に非言語的側面に焦点を当てて分析を進めている。

情報学分野における会話インタラクション研究は、転記テキストを主な対象とする第一世代、音声データを主な対象とする第二世代、ビデオデータを主な対象とする第三世代を経て、視線・動作などの計測データを対象とする第四世代に入ろうとしている。しかし、高品質・高精度な記録のための技術上の制約（例えば音声記録装置は通常2チャンネルまでしか記録できない）などもあって、二人の話者による対話に対象が限定されることが多かった。ところが、日常場面におけるインタラクションは三人以上の話者によるものが多く（例えば我々は普通三人以上からなる家族の一員として成育するし、学校・職場でも三人以上で集うことが多い）、二人対話は電話会話など限られた場面で見られない。

我々のグループでは、多人数の話者による会話インタラクションの仕組みに迫るための手始めとして、三人会話を収録し分析した。以下にその概要を述べる。

3.2 マルチモーダル三人会話の収録

日常に近い環境での会話インタラクションの収録のため、騒音の少ないラウンジに収録機器をもち込んで三人会話の収録を行った。被験者は、大学生・大学院生・OBを含む同性三人からなるグループ12組（男性・女性グループそれぞれ6組、18～33歳）で、それぞれ10分程度ずつに区切って3セッションの会話を行った。同一グループ内の被験者は友人同士で、収録に要した時間は休憩時間を含めて各40分程度である。被験者は1.5mほどの距離を隔てて、ほぼ正三角形をなすように対座し、対面で会話を行った。各会話の開始前に、被験者のいずれかがサイコロを振って会話のトピック（「情けない話」、「恋の話」、「びっくりした話」、「腹の立つ話」、「びびった話」など）を決定した。被験者は、そのトピックに固定されることなく、自由に話題を展開してよい旨を伝えられており、実際のトピックは会話中一定でない。

ラウンジに設置された4台のハンディカム（Sony DCR-VX2000）によって、各人の正面映像と全被験者を捉えた全景映像を記録した。音声は、単一指向性接話型マイクロホン（Sennheiser MKE104）を通じて、各チャンネル独立してデジタルマルチトラックレコーダ（TASCAM MX-2424）に記録した。対面であるためほかの話者の音声が多回り込むが、分析に支障のない程度であ

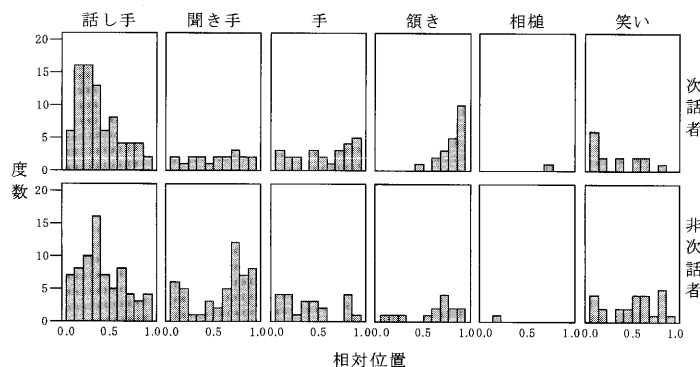


図3 聞き手単位内の相対位置に対する聞き手のちょっとした振舞いの分布。「話し手」は話し手への視線、「聞き手」はほかの聞き手への視線。[伝05]の図3・4を再構成

る。同時に、デジタルミキサ (TASCAM DM-24) でミックスダウンされた音声と、画面四分割器 (For.A MV-40E) を通して合成した映像を、DV カムコーダ (Sony DSR-1500A) に記録した。

会話データは全 36 会話 (計約 5 時間 40 分) からなり、それぞれ会話開始から 9 分 26 秒までの各被験者の正面映像、全被験者を捉えた全景映像、3 チャンネル音声、および転記テキストを含む。

3・3 三人会話における聞き手の行動の分析

会話インタラクションの仕組みの中でも、現在我々が最も関心を寄せているのは、円滑な話者交替を支えている仕組みである。二人対話では次話者として可能な参加者が一人しかいないため、話者交替の問題はタイミングの問題だけに単純化される。しかし、多人数会話では適切なタイミングに加え、誰が次話者として発話を開始するかという問題が重要になる。人々は会話の中で、誰が次話者として期待されているかということさまざまな行動を通して示しあっているはずである。このような行動を明らかにすることは、会話インタラクションの仕組みの理解につながるとともに、会話に参加するコミュニケーションロボットや実時間で被写体を切り替える会議キャプチャシステムなどへの応用が期待できる。

我々はまず女性 1 組 3 会話のデータを用いて会話中の視線行動を分析し、以下のことを明らかにした[榎本 03]。話者が交替するときは、発話末付近で話し手が次話者を見ていることが多く、次話者は話し手を見続けるか、発話末付近で話し手から視線を逸らすことが多い。つまり、話者交替時には現行話者と次話者との間で視線交差が生じる。この視線交差は、発話末より平均 2 ~ 2.5 秒程度早く始まり、次話者の発話が始まってから平均 1 ~ 1.5 秒後に終わる。このことから、話者交替は多くの場合、視線交差中に当該二人の参加者の間で起こるといえる。この結果は、二人対話に対する Kendon の分析結果[Kendon 67]と合致するものであり、話し手の視線がしばしば次話者選択手段として用いられることを示している[Sacks 74]。

一方、次話者にならない聞き手 (非次話者) は、発話末付近で話し手か次話者の一方を見続けるか、あるいは話し手から次話者に視線を移すかのいずれが多い。話し手以外の参加者に視線を向けるというのは、非次話者に特徴的なことであり、次話者にはほとんど見られない。この非次話者の視線行動は、単に話し手との視線交差の機会を放棄しているというだけでなく、他方の聞き手を次話者として推薦しているようにも見える。

次話者と非次話者との行動の違いをより詳細に見るため、我々は次に、視線以外にもあいづち・笑い・うなずき・手の動きなど会話中の聞き手のちょっとした振舞いに注目し、それらが次話者・非次話者でどのように異なるかを分析した[伝 05]。これらの分析では、女性・男性各 1 組計 6 会話から、話者交替を伴う聞き手単位 (ある発話の開始点から次発話の開始点までの区間) を抽出し、その単位中で開始された聞き手の行動を分析した。特に、0.8 ~ 2.8 秒の継続長をもつ 267 単位について、単位内の相対位置に対する各行動の生起頻度を計数したところ、以下の特徴が見られた (図 3)。

- 次話者は聞き手単位の序盤から中盤で話し手に視線を向けることが多い。
- 非次話者は聞き手単位の中盤で話し手に視線を向けることが多い。
- 非次話者は聞き手単位の終盤で他の聞き手 (次話者) に視線を向けることが多い。
- 次話者は聞き手単位の終盤でうなずきが多い。

さらに、我々は視線行動に焦点を絞り、話し手・次話者・非次話者の行動の共起関係を分析した[榎本 05a]。その結果、話し手が発話中に次話者を見ているとき、非次話者が聞き手単位の末尾付近で話し手から次話者に視線を移し替える行動が多く見られることがわかった。しかもこの行動は、それまで発話のやり取りに参加していた参加者がそこから外れる (非次話者になる) 瞬間に顕著に現れていた。一方、参与構造が変化しない (発話をやり取りするペアが変化しない) 場合には、やり取りに参加していない聞き手 (非次話者) は、ほかの二者のいずれ

かに視線を固定している。このことから、聞き手の視線行動は、次話者選択という局所的な現象のみならず、参与構造の変化という大域的な現象にも関わっていることが明らかになった。

3.4 ま と め

千葉大グループでは、多人数の話者による会話インタラクションの仕組みにせまるため、三人会話を収録し分析した。その結果、次話者になる聞き手と次話者にならない聞き手の行動、特に視線行動に顕著な違いがあることを発見し、この行動が次話者選択や参与構造の変化に関わっていることを明らかにした。これらの研究では、聞き手の役割を実際に次話者になったかならなかつたかに応じて事後的に分類したが、リアルタイムで会話に参加する人間やシステムが利用できる概念に依拠したモデル化を行うための検討を現在進めている[榎本 05b]。

4. 複数の非言語情報を利用した意図の読取り

4.1 背景と目的

近年、ロボットやエージェントの発展はめざましく、実用的な自律型エージェントが我々の生活に入ってくることも、さほど遠い未来のことではないといわれている。このような自律型エージェントを、タスクの指示や操作の習熟に時間を割くことなしに利用するには、人間と同様にエージェントと自然にコミュニケーションできることが望ましい。しかしそのためには、言語情報やジェスチャーなどの記号的なメッセージを正確に理解するだけでは十分ではない。言外に示される、あるいは隠べいしようとしても漏えいする、人間の意図を推定することが必要となる。

そこで東大グループでは、意図が漏えいする代表的な場面として「嘘」に焦点を当てた。具体的には、我々がコミュニケーション中に嘘をつくときの非言語情報の表出の特徴を調べ、複数の非言語情報による嘘の自動判別の可能性を探った。この目的を達成するため、自発的に嘘をつける自然な状況に近い実験環境を設定し、人間同士のコミュニケーションの様子を観察・分析した(実験の詳細は[Ohmoto 05a]を参照されたい)。

4.2 実験と測定装置の概要

§ 1 実験概要

コミュニケーションの中で自発的に嘘をつく環境を設定するため、インディアンポーカーをアレンジして利用した。ゲームは三人の被験者によって行われた。被験者の発話・行動は統制されておらず、自由にインタラクションすることが許された。非言語情報の表出の長期的な変化を調べるため、3か月の間隔をおき、同じ被験者に2回目の実験を行った。

§ 2 分析対象

先行研究において、嘘(欺瞞)を見破る手掛かりにな

表 1 非言語情報一覧

非言語情報	変数
視線 (3個)	発話相手を注視している割合 顔かカードを注視している割合 注視している対象が推移した割合
韻律 (9個)	音声の高さ (前半・後半・変化) 音声の大きさ (前半・後半・変化) 音声の速さ (前半・後半・変化)
表情 (1個)	目元よりも口元が早く反応したか

ると指摘されている非言語情報が多数存在する(例えば、[DePaulo 03])。このような非言語情報の中から、本実験環境において有効だと考えられるものをピックアップした。注目した非言語情報は表 1 に示すとおりである。微妙な変化には注目せず、人間にも観測が容易な範囲で抽出した。表 1 の非言語情報 13 個を独立変数とし、発話が嘘かどうかを従属変数とした。

§ 3 分析手順

二人の被験者のうち片方の被験者(被験者 B とする)は 1 回目の実験で「嘘の発話」を行わなかったため、分析対象から外した。なお、もう一人の被験者を被験者 A とする。被験者 A は 1 回目の実験の前半と後半で嘘をつく際の挙動が異なるように見えたので、これを前半と後半に分けて分析した。

分析(a): 実験 1 回目および実験 2 回目のデータを標本データとして、被験者および実験ごとにそれぞれのデータに対して判別分析を行った。

分析(b): 被験者 A の実験 1 回目の前半と後半のデータを判別分析した結果得られたそれぞれの判別式を用いて、実験 2 回目の観測データを未知データとして、判別を行った。

4.3 実験結果

分析(a)の結果を表 2 に示す。判別に対する寄与の高い変数の数は、すべてのタームで 4 個程度であった。それらの変数のみによる判別でも 75 ~ 85 % 程度の判別率を維持していた。このとき、寄与の高い変数として、視線、韻律、表情の変数が、それぞれ必ず一つは含まれていた。これは、複数の非言語情報に同時に注目する必要性を示唆している。また、判別に対する寄与の高い変数やその寄与の高さは、被験者や実験ごとに異なった。このことから、嘘をつくときの非言語情報の表出が、数十分のコミュニケーションの中でも変化することが確認された。その一方で、同一被験者内では、判別式における変数の正負が、状況にかかわらず一定である変数も多く存在した。このことは、判別に対する非言語情報の寄与の傾向は、少なくとも被験者ごとには状況によらずある程度の一貫性があることを示唆している。以上により、コミュニケーションで嘘をつく際の非言語情報の表出の特徴を、ある程度明らかにすることができたといえる。

分析(b)の結果を表 3 に示す。1 回目の前半の判別式

表2 分析 (a) の結果

被験者 A		
実験数	判別率	寄与の高い変数 (高い順)
1 回目	嘘の発話: 74% それ以外: 82%	顔かカードを注視している割合 目元よりも口元が早く動いたか 音声の大きさ (変化) 音声の高さ (前半)
	前半 嘘の発話: 79% それ以外: 83%	発話相手を注視している割合 目元よりも口元が早く動いたか 音声の高さ (変化) 音声の高さ (前半)
	後半 嘘の発話: 89% それ以外: 90%	目元よりも口元が早く動いたか 音声の大きさ (変化) 音声の高さ (前半) 顔かカードを注視している割合
2 回目	嘘の発話: 79% それ以外: 83%	顔かカードを注視している割合 目元よりも口元が早く動いたか 音声の高さ (変化) 音声の速さ (後半)
被験者 B		
実験数	判別率	寄与の高い変数 (高い順)
1 回目	—	—
2 回目	嘘の発話: 80% それ以外: 73%	目元よりも口元が早く動いたか 発話相手を注視している割合 音声の大きさ (変化) 顔かカードを注視している割合

表3 分析 (b) の結果

	嘘の発話	それ以外
前半の判別式による判別結果	36%	95%
後半の判別式による判別結果	79%	78%

で判別した場合は、「嘘の発話」の判別率が35%程度まで落ちていた。しかし、1回目の後半の判別式で判別した場合は、「嘘の発話」と「それ以外の発話」の両方で判別率は80%程度に維持された。また、実験1回目と実験2回目では3か月の期間があいていた。これらのことから、同一被験者で、ゲームの経験やプレイヤーの関係などの状況がある程度同じであれば、未知の発話に対しても非言語情報を用いた嘘の判別が可能だといえよう。

4.4 非言語情報の自動測定装置

ここまでの実験では、非言語情報の抽出を手で行っていた。しかし、エージェントへ実装するには、嘘が自動的に判別できなくてはならない。そのためには非言語情報の抽出を自動化する必要がある。音声に関しては従来の音声解析の手法を用いることができるが、視線と表情に関しては既存のシステムを同時に複数用いなくてはならない。この場合、運用上の困難が大きく、また、自然な状態に近い環境で計測することが難しくなる。そのため、ス

テレオカメラを用いた、視線と表情(顔特徴)の同時リアルタイム計測システムを作成した。作成したシステムは、目尻・目頭、上瞼、下瞼、鼻孔、口端、瞳孔中心の3次元位置と視線の方向を、20 fpsで測定することができる(システムの詳細は[Ohmoto 05b]を参照されたい)。

4.5 まとめ

ポリグラフでは接触センサによって生理指標を計測し、嘘(だます意図)を判別している。しかし、この方法では自然なコミュニケーションにおける意図を推測することが難しい。本研究では、人間が普段利用している情報から非接触で意図を推測できる可能性を示唆した。

発話の意図がわかれば、発話の一言一句を捉えなくても意味を理解できることが多い。本研究が進めば、発話の意味理解に関する知見が得られると期待される。ただし、話し手の意図は、状況や文脈からも推測されていると考えられるため、真の意味で嘘を含めた意図一般を推測するには、周囲の状況を判断することが必要不可欠となろう。

5. む す び

会話情報学研究の枠組みの中で、人間同士のコミュニケーションをそれに関わる者によって表出された行動の連鎖として捉え、かつその中で非言語情報が会話を成立させる社会的シグナルとしていかなる意味をもっているのかを明らかにしつつある研究を紹介した。紹介した三つの研究では、会話において非言語情報として表出される行動を重視しているが、(コンテンツの中の)言語情報を軽視しているわけではけっしてない。むしろ、会話情報学という観点からいえば、将来的には言語情報と非言語情報の研究は、そしてコンテンツとインタラクションの研究は統合されるべきものだと考えている。

本稿で紹介した研究成果は、会話情報学研究においていかなる意味をもつのだろうか。エージェントが人とスムーズにコミュニケーションを行うためには、人が会話において利用している認知傾向や認知バイアスをエージェントが積極的に利用する必要がある。その観点から考えると、「話し手の発話を円滑にする、話し手・聞き手間の身体的同調性」(2章)、「円滑な話者交替を支える視線行動」(3章)、「非言語情報を利用した嘘の読取り」(4章)といった研究成果は、そのまま、会話に参加するエージェントやロボット、実時間で被写体を切り替える会議キャプチャシステムなどへ応用が期待できる。特に4章で紹介した研究では、非言語情報を利用して嘘を自動的に判別するシステムの構築にも成功しており、会話情報学における工学的研究を支え得るものである。

◇ 参 考 文 献 ◇

- [Bernieri 96] Bernieri, F. J., Gillis, J. S., Davis, J. M. and Grahe, J. G.: Dyad rapport and accuracy of its judgment across situations: A lens model analysis, *Journal of Personality and Social Psychology*, Vol. 71, pp. 110-129 (1996)
- [Bernieri 88] Bernieri, F. J., Reznick, J. S. and Rosenthal, R.: Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions, *Journal of Personality and Social Psychology*, Vol. 54, pp. 243-253 (1988)
- [Charny 66] Charny, M. D.: Psychosomatic manifestations of rapport in psychotherapy, *Psychosomatic Medicine*, Vol. 28, pp. 305-315 (1966)
- [大坊 98] 大坊郁夫: しぐさのコミュニケーション—人は親しみをどう伝えあうか—, サイエンス社 (1998)
- [伝 05] 伝 康晴, 榎本美香: 聞き手のちょっとした振る舞いの相互作用について, 社会言語科学会第 15 回大会論文集, pp. 238-241 (2005)
- [DePaulo 03] DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. and Cooper, H.: Cues to deception, *Psychological Bulletin*, Vol. 129, pp. 74-118 (2003)
- [榎本 03] 榎本美香, 伝 康晴: 3 人会話における参与役割の交替に関わる非言語行動の分析, 人工知能学会研究会資料, SIG-SLUD-A301, pp. 25-30 (2003)
- [榎本 05a] 榎本美香, 伝 康晴: 聞き手たちのちょっとした振る舞いと話し手の視線, 社会言語科学会第 16 回大会論文集, pp. 240-243 (2005)
- [榎本 05b] 榎本美香, 伝 康晴, 松坂要佐: 3 人会話における談話行為と受け手のラベリングとその基礎的分析, 人工知能学会研究会資料, SIG-SLUD-A502, pp. 7-12 (2005)
- [Grice 69] Grice, H. P.: Utterer's meaning and intentions, *Philosophical Review*, Vol. 78, pp. 147-177 (1969)
- [Kendon 67] Kendon, A.: Some functions of gaze direction in social interaction, *Acta Psychologica*, Vol. 26, pp. 22-63 (1967)
- [長岡 03] 長岡千賀, 小森政嗣, 中村敏枝: 音声対話における 2 者間の相互影響—時間的側面からの検討—, 信学技報, No. 103 (113) (HCS2003-6), pp. 19-24 (2003)
- [Ohmoto 05a] Ohmoto, Y., Ueda, K. and Komatsu, T.: Reading of intentions that appear as diverse nonverbal information in face-to-face communication, *Proc. Symposium on Conversational Informatics for Supporting Social Intelligence and Interactions*, pp. 52-57 (2005)
- [Ohmoto 05b] Ohmoto, Y., Ueda, K. and Ohno, T.: The method of automatic discrimination of lies by synthetically using multimodal nonverbal information. *Proc. 3rd Conference on Computational Intelligence, Robotics and Autonomous Systems*. (2005)
- [Sacks 74] Sacks, H., Schegloff, E. A. and Jefferson, G.: A simple systematics for the organization of turn-taking for conversation, *Language*, Vol. 50, pp. 696-735 (1974)

2005 年 11 月 30 日 受理

著 者 紹 介



植田 一博 (正会員)

1988 年東京大学教養学部卒業。1993 年同大学院総合文化研究科博士課程修了。博士 (学術)。東京大学大学院総合文化研究科助手を経て、1999 年より同助教授 (2000 年 4 月～04 年 3 月同大学情報学環助教授)。専門は認知科学、認知神経科学、行動ファイナンス。日本認知科学会、情報処理学会、ヒューマンインタフェース学会、The Cognitive Science Society, AAAI 各会員。



吉川 左紀子

1977 年京都大学教育学部卒業。1982 年同大学院教育学研究科博士課程単位取得退学。追手門学院大学人間学部助教授を経て、1997 年京都大学教育学部助教授。2002 年同教授。博士 (教育学)。顔・表情・視線・身体から読み取られる情報がコミュニケーション過程でどのように利用され、意図伝達や他者理解に役立てられているのかを社会認知科学・社会認知神経科学的アプローチにより研究している。日本心理学会、日本認知科学会、日本基礎心理学会、日本認知心理学会、日本感情心理学会各会員。



伝 康晴 (正会員)

1993 年京都大学大学院工学研究科博士後期課程研究指導認定退学。ATR 音声翻訳通信研究所研究員、奈良先端科学技術大学院大学情報科学研究科助教授を経て、2000 年 10 月より千葉大学文学部助教授。博士 (工学)。専門は心理言語学・計算言語学。特に対面会話の分析・モデル化。日本認知科学会、言語処理学会、社会言語科学会、日本心理学会、日本認知心理学会各会員。



長岡 千賀

1999 年大阪大学人間科学部卒業。2004 年同大学院人間科学研究科博士後期課程修了。2004 年より日本学術振興会特別研究員 (京都大学)。博士 (人間科学)。相互作用場面におけるパラ言語・非言語行動による意図伝達に関する感性情報心理学的・認知心理学的研究に従事。日本心理学会、日本人間工学会各会員。



大本 義正 (学生会員)

2003 年東京大学大学院総合文化研究科広域科学専攻修士課程修了。現在、同博士課程在学中。人間と他者のつながりに興味をもち、マンマシンインタフェース、人のコミュニケーションなどの研究に従事。ヒューマンインタフェース学会、日本認知科学会各学生会員。



榎本 美香 (学生会員)

1998 年奈良女子大学大学院文学研究科修士課程修了。1998 年千葉大学大学院自然科学研究科博士後期課程入学。現在、同大学院在学中。2000 年 6 月～01 年 3 月 ATR 知能映像通信研究所研修研究員。専門は、認知心理学・言語心理学、特に会話におけるインタラクションの分析・モデル化。日本認知科学会、社会言語科学会、日本心理学会各会員。