

模擬遠隔聴診における相互行為パターンの マルチモーダル・マルチチャンネル連鎖分析

Multimodal and Multichannel Sequential Analysis of Interactional Patterns in a Simulated Tele-auscultation Experiment

高梨 克也¹, 堀 謙太², 内藤 知佐子³, 黒田 知宏³

Katsuya Takanashi¹, Kenta Hori², Chisako Naito³, Tomohiro Kuroda³

¹ 京都大学学術情報メディアセンター

¹ Academic Center for Computing and Media Studies, Kyoto University

² 群馬県立県民健康科学大学診療放射線学部

² School of Radiological Technology, Gunma Prefectural College of Health Sciences

³ 京都大学医学部附属病院

³ Kyoto University Hospital

Abstract: In tele-auscultation, since a doctor cannot operate an auscultator by herself, she must refer a target point by a marker and have a helper in a distant place move an auscultator by proxy. This article analyzes a simulated tele-auscultation experiment and proposes an interactional pattern observed in the process of multimodal communication from pointing by marker by a doctor, operation of an auscultator by a helper to auscultation by the doctor. This pattern is then considered in terms of division of transmission between multi-channels of a system for tele-auscultation and a tele-conference system for conversation. At the end, problems on the system environment found in the experiment are addressed.

1. はじめに

遠隔通信を利用した模擬聴診実験を対象に、その相互行為プロセスの特徴の解明を試みる。遠隔聴診では、医師が聴診器を直接操作できないため、聴診器操作を代行する介助者に指示することによって聴診が達成される。そこで、医師による指示から介助者による聴診器操作を経て、医師による聴診に至るという一連の相互行為プロセスにおける典型的なパターンを定式化する。次に、このプロセスには、聴覚と視覚という2つのモダリティと遠隔聴診支援システムとTV会議システムという2つの通信システムが介在しているため、上記の相互行為パターンについて、マルチモーダル・マルチチャンネルの使い分けの観点から考察する。さらに、今回の分析を通じて得られたシステム環境に関する課題を述べる。

2. 実験環境

堀ら[1]と同様、在宅療養等の遠隔診療において、問診・視診では医療機関への早急な搬送の必要性の判断が難しい呼吸器系疾患とした遠隔聴診による初

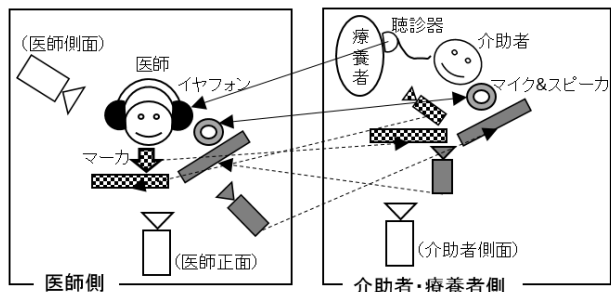
期スクリーニングを想定し、健常例と疾患例についての模擬聴診実験を行った。

聴診を行う医師と聴診対象の療養者・介助者とは別室におり、医師と介助者はTV会議システムと遠隔聴診支援システムを使用して有線LAN経由でコミュニケーションを行う。遠隔聴診支援システムでは、療養者の胸部映像に重畳表示するCGマーカの位置を医師がタッチパネル操作で指示し、聴診音を介助者は聴診器で直接、医師は通信回線経由でイヤフォンでそれぞれ聞く。被験者は聴診経験を有する医師1名と聴診経験を有しない介助者1名である。療養者役については、疾患例の設定が必要であるため、今回は聴診訓練用シミュレータを使用した。

聴診は健常例3、疾患例3の計6セッションを実施し、聴診位置は医師が自由に決めた。従って、各「セッション」の中には複数回の「位置」での聴診が含まれる。各セッション後に医師、介助者それぞれに対して、指示の正確さなどに関するアンケートを実施した。さらに、全セッション終了後に医師と介助者に個別インタビューを行った。

実験の様子は医師の正面と側面、介助者の側面に設置したビデオカメラで記録した。また、TV会議シ

システムと遠隔聴診支援システムの映像・音声もそれぞれ記録している。以降の分析では、医師側面、介助者側面のビデオカメラデータを用いる（図1）¹。



【図1】実験環境

(チェック柄は遠隔聴診支援システム，塗りつぶしはTV会議システム，白抜きは収録用機器，点線矢印は映像通信，実線矢印は音声通信を表す.)

3. 先行研究

Clark&Krych [3]は、双方向コミュニケーションにおいては、参加者が互いに対話相手の状況をリアルタイムでモニタできることが重要であるとの立場から、教示者が組立者に対して教示者の手元のモデルと同じようにレゴブロックを組み立てるように指示するという対話実験を分析している。分析の結果、この課題では、音声だけでなく、視覚的情報でのモニタリングが特に重要になること、教示者が組立者の組み立ての状況を視覚的にモニタできるため、組立者は言語的発話をしなくなることで、組立者の視覚的行為のうち、教示者は組立者が教示者にブロックを見せる場合のようなコミュニケーション的行為だけでなく、組立作業自体もモニタしながら教示を行っていることを明らかにしている。

同様に、細馬[4]も、視覚的モニタリングに着目しており、指示者からの指示に基づいて、作業者が自身のヘルメットに貼られたシールの位置を指で発見するという対話実験を行い（シールを直接指差すことは禁止）、作業者が率先して指を動かしては一時停止するという断続的行為によって、この一時停止の際に指示者が反動的に指示を発する機会を生み出す、というように、指示に従う作業者の側の能動的な視覚的行為が問題解決のためのリソースとなっていることを見出している。

¹ 収録環境の詳細については堀ら[2]を参照されたい。

² 「指差しの終了時間」は指差しが対象物に到達した時点を表しており、この指差しが対象物から撤退 retract [6]し始める時点ではない。

³ 榎本・中野[5]の焦点の一つは、人対人のコミュニケーションで観察されたこれらの連鎖規則が、教示者がヒュ

さらに、榎本・中野[5]は、パソコン画面が投影された大型スクリーンの前に立った教示者が別室にいる学習者に対して、音声通信と転送されたパソコン画面の情報を介してこのパソコンの操作（指定された位置をクリックする）を指示するという課題を行っている。具体的には、教示者による「教示発話」、「対象物への指差し」の終了時間²と学習者による「ポインタの対象物への滞留」、「マウスクリック」の開始時間の間のタイミングの分析に基づき、次の2つの連鎖規則を発見している³。

- ・教示者の指差しが終了するところで、学習者はマウスを対象物の上へ置く。
- ・教示者による指差しに伴う発話が終了するところで、学習者はマウスクリックを行う。

特にこの研究は、教示者の指差しという視覚的行為による位置指示と学習者によるこの位置へのマウスポインタの移動とクリックという視覚的応答が課題の中心となっていることと、通信システムを介した遠隔コミュニケーションを対象としていることの2点において、本研究の課題と類似しているため、次節ではこの知見を参照した分析を行う。

4. 聴診のマルチモダリティ連鎖

今回の実験では、人間の療養者ではなく聴診シミュレータを用いたこともあり、聴診のサイクルはかなり規則的かつ単純なものであったため、典型的な「聴診サイクル」を次のようにパターン化できる。以下、特徴的なステップとステップ間の行為連鎖について順に考察する。

1. 医師：マーカー指示⁴
2. 介助者：聴診器移動の開始
3. 医師：言語的指示「(ここ/こちら)お願いします」(省略可)
4. 介助者：言語的応答「はい」(省略可)
5. 介助者：聴診器位置の確定(圧着)
6. 医師：聴診と記録
7. 医師：次のマーカー指示(1に戻る)(図2)
8. 介助者：聴診器圧着終了と次の位置への移動開始(2に戻る)

マンエージェントに代わった時にどのように変化するかという点にもあるが、本稿では人対人の結果との比較のみを行う。

⁴ 介護者による聴診器移動とは異なり、マーカーによる指示では、結果としてマークされた点のみが表示されるため、移動中の軌跡などの情報は利用できない。



[図2] 遠隔聴診支援システム画面
(介助者が聴診器を圧着させている最中に医師が次の位置をマークした瞬間。ステップ7に相当)

4.1 指示－受諾連鎖とマルチモダリティ

医師によるマーカ指示（上記1）と介助者による聴診器移動の開始の間には前者が達成されないと後者が実行できないという論理的依存関係がある⁵。しかし、新たな位置の聴診開始時に、医師はマーカによって新たな位置をマークするだけでなく、直後に「(ここ/こちら) お願いします」という言語的指示も発する（3）ことが多い（両者の順序については医師へのインタビューでも確認され、データ上でも基本的に不変であった）。

そこで、介助者による聴診器移動（2）が言語（3）と非言語（1）のどちらに対する応答なのかを特定するため、介助者側の映像に基づいて2と3の前後関係を調べた（表1）⁶。聴診器移動は指示に対する介助者による「応答」であるため、応答者側の映像に基づくのが妥当である。ここから分かるように、介助者は医師発話の開始前に聴診器移動を開始していることが圧倒的に多く、聴診器の移動は医師によるマーカ指示への「受諾」の応答と見なすべきものである。実際、介助者はある位置への聴診器の圧着を完了した後は遠隔聴診支援システムの画面を注視し続けることによって、次のマーカ指示にすぐに対応できるよう備えている（後述）。

さらに、医師によるマーカ指示はこの課題において省略不可能なのに対して、言語的指示の方は後半のセッションになるにつれて省略される場合も出てくる（表1「うち医師発話なし」）。医師へのインタビューでも、「お願いします」と言わないのは「映

像上で（介助者の手が）そちらに向かって動いている」のが確認できるからであるとの回答を得ている。このように、介助者が医師によるマーカ指示を視覚的にモニタすることによって聴診器移動を開始しているだけでなく、医師の側でも聴診器移動を視覚的にモニタすることによって、自身の視覚的指示の基盤化を確認しており、その結果、場合によっては言語的指示を省略する場合もある⁷。また、介助者による言語的応答「はい」は必ず医師の言語的指示発話の末尾に合わせて生起しており、言語的指示が省略されると「はい」も省略される。

以上の観察は次のようにまとめられる。

・モダリティ内連鎖の優先性

医師によるマーカ指示（1）に対しては介助者による聴診器移動の開始（2）が、また、医師の言語的指示（3）に対しては介助者の言語的応答

（4）がそれぞれ対応するというように、二者の間の行為連鎖は視覚的行為同士、言語的行為同士の間でそれぞれ形成される。

・視覚的連鎖の優先性

視覚的行為連鎖は言語的連鎖よりも時間的にやや先行して行われ、省略不可能であるのに対して、言語的連鎖の方は省略可能である。

前者は榎本・中野[5]で見出された2つの「連鎖規則」（第3節）と一致する（榎本・中野(2008)では、言語的応答でなくマウスクリック）。ただし、「聴診器移動」は榎本・中野[5]の「(ポインタの)対象物への滞留」に対応するものの、「対象物への滞留」はポインタが画面上の対象物の範囲内に入っている時間区間のみを指しているのに対して、本研究では、聴診器の対象位置への到達（5.聴診器位置の確定）だけでなく、その移動の「開始」（ジェスチャーフェーズにおける準備 preparation[6]に相当）に着目しており、これにより、上記の2種類の連鎖関係がより明確に整理できるようになった⁸。

なお、7と8は、同じセッションの中での2回目以降の位置における1と2と重複するが、これらを冗長に明示したのは、介助者にとっては「聴診器をいつまでその位置に圧着しておくべきか」という問題が存在するからである。介助者にインタビュー

⁵ 1と2の間の前後関係を定めるもう一つの要因については本4.1節の最後で述べる。

⁶ 「医師発話なし」と「介助者「はい」なし」、「追加位置」と「前/背面最初」はそれぞれ重複可。「追加位置」は前面と背面の基本6位置以外の位置を指し、「前/背面最初」は前/背面での連続する複数位置のうちの最初のもので、介助者はこの位置（の直前）でシミュレータを回転させたり聴診器を持ち直したりしていることが多い。

⁷ この課題ではマーカの指示位置が明確であるため、介

助者が聴診器移動を一時中断させて医師からの指示を誘発するといった、細馬[4]で見られたようなパターン（第3節）は見られなかった。

⁸ 聴診器（あるいはこれを操作する介助者の手）の動きに比べて、榎本・中野[5]のポインタ移動は明確に観察するのが難しいという事情もあると考えられる。インタラクション相手の準備動作などをモニタできることが重要な課題の場合には、ポインタの移動などを相手にとってもより分かりやすく表示することが望ましい。

[表 1] 介助者の聴診器移動開始と医師発話の生起順 (介助者側映像より)

セッション	位置数	(うち追加位置)	聴診器移動開始→医師発話	(うち医師発話なし)	(うち介助者「はい」なし)	医師発話→聴診器移動開始	(うち追加位置)	(うち前/背面最初)
1	16	(4)	13	(0)	(1)	3	(2)	(1)
2	20	(8)	17	(1)	(0)	3	(1)	(2)
3	24	(12)	24	(4)	(4)	0	(0)	(0)
4	16	(4)	16	(1)	(1)	0	(0)	(0)
5	21	(9)	20	(1)	(1)	1	(0)	(0)
6	12	(0)	12	(5)	(5)	0	(0)	(0)
計	109	(37)	102	(12)	(12)	7	3	3

したところ、この点について医師からの言語的な指示がなくても、医師のマークが「別のところを指した」ら前の位置での聴診器の圧着を終了してもよいことが分かるという回答が得られた。このように、この課題においては、ある位置の聴診の「開始」時に1.マーク指示と2.聴診器移動の間に視覚モダリティでの基盤化[7]が見られるだけでなく、当該位置での聴診「終了」についても、7. 医師による次のマーク指示と8. 介助者による聴診器圧着終了の間で視覚モダリティ上での基盤化が達成されている。

4.2 通信遅延による問題の複雑化

しかし、2と3の間には時間的な生起順に関する規則性があるだけで、1と2、3と4の間のような連鎖関係はない。この点について興味深いのは、指示に回答する介助者の視点からは上記のパターンが認定されるものの、医師側の映像では、医師は2の聴診器移動が遠隔聴診システム画面で観察されるより前に3の言語的指示を始めている場合も多いという点である。これはこのシステムの通信遅延による(図3)。このように、介助者の応答タイミングに基づくと、上記のようなモダリティ内連鎖の優先性が明らかになるものの、医師の側からは自身の言語+非言語での複合的な指示に対して介助者が言語+非言語で複合的に応答しているように見えている可能性もある。今回は介助者と医師の間でのこうした認識のずれに起因する明らかな問題は見出されなかったが、十分に注意する必要がある。

4.3 聴診のマルチモダリティ

協同活動の目的が医師による「聴診」(聴診音を聞き症状についての判断を下す)(上記6)にあるという点は、Clark&Krych[3]や榎本・中野[5]と異なる本課題の特色である。これらの研究では、言語的

発話と視覚的行為・状態という2つのモダリティだけが関与していたのに対して、本課題では、「聴診音」という外部観察者にはアクセスできないモダリティこそが課題の成否を分ける核となる。

この外部観察困難なモダリティについて、医師のインタビューから、「(聴診中は)目を瞑っていることも多い」「(聴診音を)聞く方に集中していると(画面は)見えていないはず」という貴重な証言を得た。これは聴診時には視覚から聴覚へと意識が移動するということであると考えられる。「聴診」はTV会議システムでは無音区間になるため、外部観察的に言語情報と視覚情報だけを分析する立場ではこうした観点が見失われやすいことに注意が必要である⁹。

5. マルチチャネルシステム環境

遠隔支援聴診システムとTV会議システムという2つの通信システムは音声・映像チャンネルでの情報の送受信を分担している(表2,表3)。カッコ内の数字は前節の相互行為パターンでのステップ番号に対応している(1,2と重複する7,8は省略)。

[表 2] 遠隔聴診支援システムのチャンネル

	音響	映像
医師→介助者	(なし)	マーク位置(1)
介助者→医師	聴診音(6)	聴診器位置(2,5)

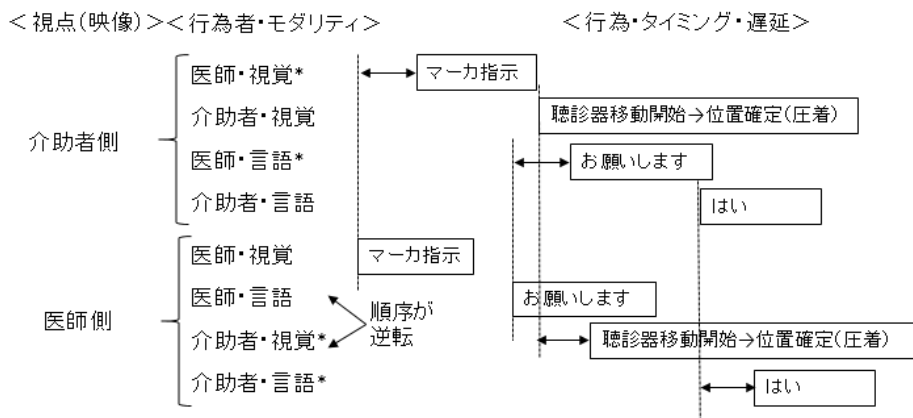
[表 3] TV会議システムのチャンネル

	音響	映像
医師→介助者	指示(3)	不要?
介助者→医師	受諾(4)	聴診器位置は今回は使わなかった

行為連鎖1-2は同じシステムの映像チャンネル同士、3-4は同じシステムの音声チャンネル同士という同一システム・モダリティ間での連鎖であるのに対して、5-6では映像→音声という異モダリティ間での連鎖が形成されているという特徴がある。

⁹ 聴診フェーズにおいては、聴診結果としての所見を記録する作業も重要になる。しかし、メモは各位置の聴診終了後に必ず記入されていたわけではなく、今回はその

相互行為連鎖上での役割に関する規則性は見いだせなかった。聴診音の違いと関連している可能性は考えられるが、この点は今後の課題としたい。



[図3] 介助者側/医師側で知覚される行為順・タイミングのずれ (*は遠隔での行為, ←→は通信遅延)

また、今回の遠隔聴診実験では、TV 会議システムの映像チャンネルはほとんど使われておらず、また、前節で述べたように、言語的指示と受諾のやりとりは省略されることもあることを考慮するならば、やりとりの大半は遠隔聴診支援システムの映像・音声チャンネルを通じて行われていたといえる。しかし、これはTV 会議システムに担われているチャンネルが不要であるということは意味しない。

まず、音声チャンネルについて、各セッションの最中に少なくとも1回は療育者(シミュレータ)の身体の前後を回転させる必要があるが、この指示はTV 会議システムの音声チャンネルを通じて行われており、介助者もこの指示を聞いてからでないと行動できないという意味で、このやりとりはTV 会議システムの音声チャンネルで主導されているといえる。今回の実験ではシミュレータを使っており、療養者の身体に関する操作は単純化されていたが、実際の人体を対象とする場合には、姿勢などに関するより詳細な指示が必要になることも考えられる。その他にも、回数は少ないものの、医師が介助者に対して、聴診器の当て方に関する修正指示(3回)や聴診器の電源(一定時間で自動的にOFFになる)を再度ONにするよう指示する(1回)ケースも見られた。Clark[7]はメインのやりとりにトラブルが生じた際にはこれを解消するための別の「トラック」でのやりとりが必要になると指摘しているが、遠隔聴診のような通信システムを介したコミュニケーションでも、パターン化されたもの以外のやりとりが必要になった際には、TV 会議システムの音声チャンネルのような別の「トラック」が確保されていることが重要になる¹⁰。

¹⁰ さらに、この課題で最も重要なことは医師が聴診音を正確に聞き取ることであるため(4.3節)、イヤホン経由での聴診音に音声会話の情報(現在はイヤホンではなくスピーカーで再生)が混入するのも望ましくない。

次に映像について、今回の実験では、聴診器位置は遠隔聴診支援システム画面とTV 会議システム画面の両方でモニタできる環境にあったが、インタビューによれば、医師は課題遂行中は自身が指示したマーカ位置も確認できる前者の画面のみを参照していた。しかし、これは今回の実験がシミュレータを用いた、かなりルーティン化したものであったためであるとも考えられる。Clark&Krych[3]が指摘するように、課題が順調に進行している間是对話相手のマーカや聴診器などの操作状況のモニタリングの方が重要だが(本課題では遠隔聴診支援システムが担当)、トラブル発生時には対話相手の顔が視覚的モニタリングの焦点となることも予想される。従って、今後より多くの症例に対する聴診や、聴診以外に遠隔診断システムの開発を進めていく際に、医師と介助者の間での視覚情報を含めた自由な会話が可能な環境が不要であると即断することはできない。

6. システム環境の評価

第4節で考察したように、今回の実験では、聴診の相互行為はかなり単純なパターン化されたものであったが、診断結果は正常音条件のすべてで外れていた。その原因としては、医師が「(症状が)あるものと思って診察していた」という点もあるが¹¹、特に映像画面については、システム環境上の問題も見られた。

まず、遠隔聴診支援システムのカメラアングルの関係でカメラから遠い画面右側が見えにくいという問題があったと考えられる。第5節でも述べたように、6セッション中3回と事例数は少なかったもの

¹¹ 今回想定した場面が初期スクリーニングであることを考えれば(第2節)、疾患例を正常と判断する場合ほどのリスクはないとも言える。

の、聴診器の当て方について医師が明示的に修正の指示をするケースが見られたが、これらはいずれもシミュレータの向かって右下で起こっており、セッションごとのアンケートでも「右下の聴診時に圧着が弱く、聞こえにくい」という回答が複数回あった。今回の遠隔聴診システムでは、1. 療養者の胸部とそこでの聴診器操作の様子を含む映像が介助者側から医師側に送られ、2. 医師がこの映像に対してタッチパネルで指示したマーカ位置が追加された同動画が介助者側に返送され、3. 介助者はこの映像を自分の側の画面で確認し指示されたマーカ位置に聴診器を圧着させる、というプロセスを経るため（図1、図2参照）、1の段階でのカメラアングルの問題は医師が聴診器位置を確認する際と介助者がマーカ位置を確認する際の両方で問題を引き起こす。

次に、遠隔聴診支援システム画面の縦横比の問題も見られた。実験後に両画面を比較してみると、TV会議システムの画面の方が聴診器が下に当たっているように見えることが判明した。これは遠隔聴診支援システムのオリジナルのカメラ映像が縦横比4:3であったのに対して、医師側のディスプレイは16:9で、オリジナルよりも圧縮されて表示されていたことによる。この点については、医師へのインタビューでも、「右の下はこの向きでは画像が圧縮されている感じで、少しカーブの具合も分からない」と指摘されており、画面右下では上述のカメラアングルの問題と縦横比圧縮の問題が同時に生じていたとも考えられる。

さらに、この点に関連して、療養者の身体という3次元情報を2次元画像から理解することの難しさもある。事後インタビューでは、「曲がり方とか体の奥行きが画像で私が見えてないので、指定している位置が思っているのと少しズレがあるのかもしれないと思った」や、例えば肩甲骨の曲がり角のあたりについては介助者が聴診器を圧着するのが難しいと予想されたため、もう少し平らなところにマーカ位置を変更するというケースがあったと報告されている。遠隔聴診支援システムのカメラアングルと画面の縦横比についてはすぐに改善可能であるものの、3次元の奥行きの問題についての本格的対処は困難であるため、まずはどのような場面でどのような困難が生じるかを引き続き検討することが必要になると考えられる。

7. まとめと課題

本研究では、医師が聴診器を直接操作できず、遠隔の介助者が聴診器操作を代行する遠隔聴診に見られる典型的な相互行為パターンを特定し、これにつ

いて、聴覚と視覚という複数のモダリティと遠隔聴診支援・TV会議という2つの通信システムのマルチチャンネルの間での情報の使い分けの観点から考察するとともに、今回観察されたシステム環境に関する問題点をまとめた。

今回は療養者役としてシミュレータを用いたが、今後は、人間を対象とした場合に、身体姿勢などに関するより具体的な指示が生じることによって相互行為パターンがどのように複雑化し、システム環境についてのどのような課題が生じるかについても検討したいと考えている。

謝辞

本研究の一部は、平成25年度国立情報学研究所共同研究「MR技術を用いた遠隔聴診支援におけるコミュニケーション分析」（代表者：黒田知宏）、平成26-28年度科研費補助金基盤研究（B）「発話連鎖アノテーションに基づく対話過程のモデル化」（代表者：傳康晴）の助成による。

参考文献

- [1] 堀謙太, 内田祐介, 南麻弥, 他: 遠隔聴診の進行への聴診器操作教示システムの影響, VR医学, 12(1), 15-26, 2014.
- [2] 堀謙太, 高梨克也, 内藤知佐子, 黒田知宏: 模擬遠隔聴診におけるコミュニケーション分析の試み, 第19回日本医療情報学会春季大会抄録, 2015.6 刊行予定
- [3] Clark, H. H. & Krych, M.: Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50: 62-81. 2004.
- [4] 細馬宏通: 対話における応答詞「あ」の機能: 話し手にとっての「新規情報」は相互行為にどう利用されるか? 社会言語科学会第16回大会発表論文集, 32-35. 2005.
- [5] 榎本美香, 中野有紀子: 人-人, 人-ヒューマンエージェントの社会的インタラクションにおける言語・非言語行為の配置規則. 知能と情報, 20 (4): 540-556. 2008.
- [6] Kendon, A.: *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- [7] Clark, H. H.: *Using language*. Cambridge University Press. 1996.