ビデオ対話における相互作用過程の分析¹ ——対象指示コミュニケーション課題による検討——

筑波大学 南部美砂子^{2,3}・法政大学 原田悦子

Interactive processes in video-mediated referential communication

Misako Nambu (*Research Fellow of the Japan Society for the Promotion of Science, Doctoral Program in Psychology, University of Tsukuba, Tennodai, Tsukuba 305-8572*) and Etsuko T. Harada (*Faculty of Social Sciences, Hosei University, Aihara-machi, Machida 194-0298*)

This study examined interactive processes that occurred in video-mediated referential communication task. Pairs of participants were assigned to one of four conditions; in a room with partition, or in separate rooms linked by telephone with three levels of noise interference (none, low, and high). In all conditions, a video-monitor link allowed the participants to see the upper body of their partner. Analyses of gestural and visual interaction revealed that participants in partition and no-interference conditions did not rely on the visual channel, while those in the others frequently looked at the monitor. Furthermore, gestural references appeared especially in high-noise condition. Qualitative analyses of conversations showed that participants in high-noise condition relied more on demonstrative references, and used more metaphoric expressions of a whole referent, based on the common ground shared by the pair. These results are discussed in terms of how people collaboratively interact and construct a shared context for communication between them.

Key words: video-mediated communication, referential communication task, gestural and visual interaction, common ground, collaboration.

近年、通信技術の発展やネットワークの整備にともない、遠隔地間の対話を支援する新たな対話システムが次々に登場し、対話の形式が多様化してきている。その中でも特に、テレビ電話やビデオ会議システムに代表される音声と映像が利用可能な“ビデオ対話システム”は、さまざまな領域から期待と注目を集めている (e.g. 柿井, 1997)。

ビデオ対話に関するこれまでの研究の多くは、映像チャンネルの有無に始まり、映像・音声の品質や大きさ、映像と音声の同期 (転送遅延)、音声の伝達形式 (全二重/半二重)⁴、アイコンタクトや視線方向のサポートなど、対話システムのさまざまな属性を実験要因として取り上げ、これらが対話課題の成績や効率 (試

行時間や発話量)、話者交替などの構造的側面に及ぼす影響について検討してきている (レビューとして Finn, 1997)。これらの研究に共通する特徴として、第1に、対面对話を理想的な対話形式と位置づけ基準としている点、第2に、対面对話を目標とする技術的可能性に研究そのものが駆動されている点が挙げられる。そのため多くの研究が、“対話システムにおいて利用可能なモダリティの数が多いほど対話の質が向上する”というモダリティ還元主義 (川浦, 1993) 的な仮説や、“音声や映像の品質、システムの性能を向上させるほど対面对話に近づく”という技術決定論的な仮説を検証する形で進められている (詳しくは、南部・原田・海保, 2000)。

しかしながら実際には、このような仮説の妥当性を示す知見はほとんど得られていないだけでなく、このような枠組みではとらえきれない事例や現象も数多く報告されている。例えば、ビデオ対話システムは電話やテキスト対話システム (e.g. 電子メール、チャット) と比較して主観的・感性的評価が低く嫌われがちであることや、映像チャンネルが実際にはほとんど利用されないこと (原田, 1997)、映像チャンネルを介

¹ 本研究は筑波大学心理学研究科に提出した修士論文の一部を加筆・修正したものである。本研究の一部は、1998年日本認知科学会第15回大会 (南部・原田, 1998b)、1999年 The 2nd International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference (Nambu & Harada, 1999) において発表された。

² 現所属：日本学術振興会・法政大学

³ 本論文の執筆にあたり、筑波大学心理学系の海保博之先生にご指導いただきました。また、査読者の方々、ならびに筑波大学心理学系 (現所属：駿河台大学) の青山征彦氏、日本学術振興会・筑波大学心理学系 (現所属：東京工業大学) の岩男征樹氏より有益なコメントをいただきました。心より感謝いたします。

⁴ 全二重 (full-duplex) では電話のように複数の対話者が同時に話すことができるが、半二重 (half-duplex) ではトランシーバーのように一度に一方のみに制限される。

した相互作用では対話上の問題が生じやすいこと (Heath & Luff, 1991) などである。したがって、ビデオ対話システムを今後さらにさまざまな分野で活用することを視野に入れ、ビデオ対話に関する理解を十分に深めるためには、これまでの研究アプローチでは不十分であると言わざるを得ないだろう。

このような背景から、本研究では、ビデオ対話における対話者の認知活動や対話者間の協働 (collaboration) に焦点を当てた新たなアプローチを試みる。

具体的には、原田 (1997) が人工物研究の立場から提案する“人工物間比較”に基づく検討を行う。人工物間比較とは、特定の課題において複数の人工物が使われる様子を観察・比較し、その相違から、人工物の使用者自身の視点から見た認知的課題を解明する方法である。本研究では、実験における対話課題として対象指示コミュニケーション課題 (referential communication task) を利用し、対話条件としてビデオ対話システムの有無に加え、その音声チャンネルに雑音による障害を発生させた条件を設定する。

まず、障害状況をも含めた対話条件の設定は、次の意図によるものである。電話対話を取り上げた南部・原田 (1998a) では、実際にも起こりうるような障害のある対話環境 (e.g. 携帯電話における電波障害) を想定し、音声に雑音を加えた条件を含めた検討を行った。そこでは、障害の有無や程度に応じて対話者間の言語的協働の様態が変化することが見出されている (詳しくは後述)。これを受けて本研究では、一般的なビデオ対話システムでは映像チャンネルがほとんど利用されないという前述の報告 (原田, 1997) も踏まえ、ビデオ対話において音声に障害が発生している対話条件を設定して、障害状況下では映像チャンネルが利用されるか否か、またどのような形で利用されるか、さらにはその際どのような言語的相互作用が行われるかを分析する。このような試みを通して、ビデオ対話における対話者の認知活動や協働の特性を明らかにしていくのが本研究の目的である。

ところで、本研究が対話課題として利用する対象指示コミュニケーション課題は、特定の対象の指示・同定を目的とする課題であり、これまでに言語心理学を中心としてさまざまな研究が行われてきている。例えば、Fussell & Krauss (1989a, 1989b) は、無意味線画を用いた対象指示コミュニケーション課題 (八つの線画の並び順に関する指示メッセージの作成) において、メッセージの受け手との間で共有されている情報、すなわち共通基盤 (common ground) が、大きいと推測される場合には比喻による簡潔な表現 (e.g. 歯ブラシのよう) が多く用いられ、小さいと推測される場合には図形記述的で冗長な表現 (e.g. 縦の線が5本あり、その下が長い横の線につながっている) が用いられることを示した。また、比喻的な指示により対

象が正しく同定されるか否かは対話者が実際にどのような共通基盤を有しているかに依存すること、図形記述的な指示は一般的他者との共通基盤に基づくため比較的理解されやすいことが明らかにされている。また Clark & Wilkes-Gibbs (1986) は、タングラム図形 (後述) を用いた対象指示コミュニケーション課題 (ペアによる12図形の並び順の指示・同定) における言語的相互作用の分析から、対話者は互いに相手からの応答を情報源として即時的に共通基盤の推測・更新を行っており、その結果として、同一の指示対象に関する共有理解をより効率的に達成できるようになることを明らかにしている。このように、対話相手との共通基盤を活用し共有理解を達成することは、対話の内容や目的を超えて対話者が行うべき必須の、基本的な認知的課題であると言える。そこで本研究では、対象指示コミュニケーション課題を利用することによって、このような対話者の認知活動や言語的協働の様態が、対話システムのもたらす状況に応じていかに変化するかを明らかにしていく。

相互作用過程の分析は以下の二つの観点から行う。一つは、“映像チャンネルの利用行動”である。これまでの研究の多くは、映像チャンネルが利用可能であることと実際にそれが利用されることを同一視する傾向があり、映像チャンネルが利用されているか否か、またどのような形で利用されているかについてはほとんど問題にしていなかった。これに対し、エスノメソロジーの立場からビデオ対話の分析を行った Heath & Luff (1991) は、対話者が話し始める前に相手を見つめたり手を振ったりして合図を送る様子を描き出し、さらにこのような視覚的アクセスの多くが対話相手に見過ごされたりその意味が十全に理解されないなどの問題を生じていることを見出した。このことから彼らは、ビデオ対話において相手を“見る”ことは直ちに可能だが、相手に“見られている”のに気づくことや、自分の身体を効果的に“見せる”ことは困難であり、映像チャンネルを介すことによって視覚情報のもつ影響力が減じられる点を指摘している。

さらに、映像チャンネルが媒介する情報について、原田 (1997) は、一方にとつての“視覚情報”が他方にとっては“身体情報”であり、正確には“身体-視覚情報”としての性質をもつことを指摘する。すなわち、音声チャンネルではすべての対話者が同じ情報を受け取ることができるのに対し、映像チャンネルの身体-視覚情報は対話者にとって“非対称性 (Farr, 1991)”の強い対話資源となっており、そのことが映像チャンネルの利用行動に影響を及ぼすと考えられている。これらの点を踏まえ、本研究では、ビデオデータから映像チャンネルの利用行動を分析し、身体-視覚的な相互作用の様態が、対話システムの物理的特性に応じていかに変化するかを検討する。

もう一つの分析対象は“指示発話”である。対象指示コミュニケーション課題を用いて電話対話の検討を行った南部・原田(1998a)は、同室でのついでに越すと、別室で電話を用いる対話条件とともに、電話の音声に高/低レベルの雑音を加える条件も含めて発話内容の比較を行った。その結果、指示のはじめの発話において、前者2条件では図形記述的な指示表現が多く用いられたのに対し、音声に障害のある後者2条件では比喩による指示が多く示された。このことから南部らは、電話という音声対話システムの物理的特性に応じて、共通基盤を活用し共有理解を達成するという協働的な相互作用の様態が大きく変化することを見出した。そこで本研究では、南部・原田(1998a)が用いた対象指示コミュニケーション課題と分析枠組みを利用して、ビデオ対話における指示発話の特徴が、対話システムの特性に応じていかに変化するかを明らかにする。

方 法

実験計画 2要因混合計画。被験者間要因として対話条件4水準(ついでに+モニタ/テレビ電話/低雑音テレビ電話/高雑音テレビ電話)、被験者内要因として課題試行6回を設定した。

実験参加者 大学生1,2年生96名が、48組の友人同士のペアとして参加した⁵。各ペアは来室順に四つの対話条件に割り当てられた。

課題 Clark & Wilkes-Gibbs(1986)の対象指示コミュニケーション課題を参考に、予備調査の結果に基づき、“人の形に見える”という点で類似性が高く、“指示のしやすさ”の評定が適度にばらついている8種類のタングラム図形を選択した(Figure 1)。各図形はA6サイズのカードに一つずつ印刷され⁶、8枚のカードセットが2組用意された。図形の並び順は1回

Figure 1. 対象指示コミュニケーション課題において指示対象として用いたタングラム図形。

⁵ ついでに+モニタ条件11組、テレビ電話条件13組、高/低の雑音テレビ電話条件各12組であった。

⁶ 各カードには小さく“上”と“下”を記入した。

ごとに乱数表によって決定した。

装置 高/低の雑音テレビ電話条件では、音声にピンクノイズが加えられた(SONY BGM & RECセクター TL-12を使用)。その大きさは、予備実験において任意に設定した5水準の中で、“雑音は気になるが、会話にはほとんど支障がない”と評定された水準を低雑音、“雑音は非常に気になり、かなり会話づらい”と評定された水準を高雑音として選択した。各対話者が使用する机には、カードを並べるためのフレームを描いた紙を貼り、その右側に10型のテレビモニタ、その後方にビデオカメラを設置した⁷。ついでに+モニタ条件では机を同室で向かい合わせに配置し、その間に不透明のついでにを設置した。テレビ電話を利用する3条件では、各部屋にプッシュ式の電話機を設置し、通話には内線用の回線を利用した。音声を記録するためピンマイクを用いた。

手続き 実験はペアごとに個別に行った。ペアはくじ引きにより指示者役か行為者役に割り当てられた。課題は同一のカードセットを互いに相手のカードが見えない状況で正しい順序に並べることでであると教示し、指示者には正解となる並び順でカードがフレームに提示され、行為者にはランダムな順序に束ねたカードセットが渡された。話し合いは互いに十分であると思われるまで行ってよいとしたが、その際、図形の指示・同定は左側から一つずつ順を追って行い、順序をとばして後回しにはしないこと、またカードを直接ビデオカメラに向けて指示・確認をしないことを、ルールとして教示した。

教示後、ペアは、ついでに+モニタ条件では同室でついでにをはさみ向かい合う形で、テレビ電話を利用する3条件ではそれぞれ別室に移動して、用意された机の前に座った。両者の準備が整ったところで実験者は指示者のフレームにカードを並べ、ついでに+モニタ条件では実験者の合図によって、テレビ電話を利用する条件では指示者が行為者に電話をかけることによって、課題が開始された。課題の開始・終了時には各自でモニタの電源を操作した。

1試行終了するごとに質問紙への回答を求め⁸、両者とも記入を終了したところで再び実験者が前回とは異なる順序でカードを並べて次の試行が開始された。すべての試行と質問紙への回答が終了するまで、課題の成否や正しく同定された図形数などの情報は一切与えなかった。終了後、ペアは再び対面し、自省報告を求められた。実験全体はおよそ30分で終了した。

⁷ この設定では、対話者の視線方向は一致するが高さが若干ずれているために、完全なアイコンタクトはとれなかった。

⁸ 質問紙の内容および分析結果については南部・原田(1998c)を参照のこと。

結 果

課題試行中の発話および行動は記録され、発話はすべて書き起こされた。課題の誤答率は全体で4.2%であり⁹、第1試行で若干多かったが第2試行以降はほとんど見られなかった。

映像チャンネルの利用行動

ビデオデータによると、対話者の視線は主として手元の図形に向けられていたが、指示者・行為者ともに時折顔を上げてモニタを一瞥する行動が見られたため、これを“モニタ注視”として生起数をカウントした。さらに、モニタやカメラに向けて空中に指で図形の輪郭を描く (Figure 2)、上半身で図形の様子をまねるといった、自らの身体を見せる行動が見出された。このような“身体的指示”は、例えば“右側が三角”や“こういうやつ”のような指示発話ユニット (後述) とともに生起しており、その対応を単位としてカウントを行った。

まず、モニタ注視の生起数について対話条件と試行、役割 (指示者/行為者) の3要因による分散分析を行ったところ、条件と試行の主効果、および両要因の交互作用が有意であった (対話条件, $F_{(3,82)}=38.61$; 試行, $F_{(5,410)}=47.30$; 対話条件×試行, $F_{(15,410)}=9.03$, いずれも $p<.01$)。役割による差異は示されなかった。下位検定¹⁰の結果、Figure 3に示すように、ほとんどの試行で高/低の雑音テレビ電話条件のモニタ注視が多く、またテレビ電話を用いた3条件では特に前半の試行で試行ごとに減少する傾向が示された。

このことから、同室のついたて越しでモニタが利用可能な条件と、別室で通常のテレビ電話を用いた条件

Figure 2. 図形の輪郭を指先で示す身体的指示。

⁹ ほとんど誤答が生じていなかったテレビ電話条件を除く3条件を取り上げ、第1試行から第3試行を対象として正しく同定された図形数について分散分析を行ったが、対話条件による差異は示されなかった。

¹⁰ 多重比較の検定はすべてLSD法を用いて行った。

では、課題における役割の違いにかかわらず、1図形につき1回程度しかモニタに視線を向けておらず、音声情報により依存的な対話を行っていたことが明らかになった。その一方で、音声に障害のあるテレビ電話を用いた場合には、役割の違いにかかわらず、特に第1試行においてより積極的に映像チャンネルへのアクセスが行われていた。

続いて、身体的指示については、高雑音テレビ電話条件以外の3条件ではほとんど生起していないことが明らかになった (1試行当りの平均、ついで+モニタ0.55回、テレビ電話0.35回、低雑音テレビ電話0.19回、高雑音テレビ電話6.79回)。また、身体的指示の生起はタングラム図形の指示という課題の性質、およびそこでの指示者・行為者という役割に依存し、特に指示者において増大すると考えられる。そこで、高雑音テレビ電話条件における身体的指示の生起数について、指示者・行為者ごとに分析を行った。それぞれ試行を要因とする分散分析を行った結果、指示者・行為者ともに前半の試行で試行ごとに減少する傾向が示された (指示者, $F_{(5,55)}=3.75$, $p<.01$; 行為者, $F_{(5,55)}=4.30$, $p<.01$)。

このことから、特に高雑音状況下の第1試行では、指示者・行為者ともに映像チャンネルを活用し、身体・視覚的な相互作用を積極的に行っていたことが明らかになった。

指示発話

指示者の発話量について、文節を単位としてカウントし、対話条件と試行の2要因による分散分析を行ったところ、試行の主効果が有意であった ($F_{(5,220)}=121.68$, $p<.01$)。対話条件による差異は示されなかった。下位検定の結果、Figure 4に示すように、いずれの条件でも特に前半の試行で試行ごとに発話量が減少する傾向が示された。

Figure 3. モニタ注視の平均生起数。

Table 1
指示カテゴリー

カテゴリー名	内 容	例
ラベル	命名による指示, またはそれに準ずるもの	バレリーナ/フラミンゴっぽいやつ
全体的メタファ	図形の全体的なイメージ・動作についての比喩的な指示	片足で立っている/スケートしている人
部分的メタファ	図形の構成要素についての比喩的な指示	手は左側に2本あって/足が後ろに伸びている
図形的描写	図形の構成要素についての幾何学的・図形的描写	視覚がやや右よりの上にあって/左側に三角形がとび出ているやつ
指示代名詞	指示ユニットを構成する指示代名詞	こういうやつ/こうなっている

これまでに見てきたように、映像チャンネルの利用行動の生起ならびに指示発話の量的側面は、各図形について初めて指示・同定を行う第1試行において顕著であるため、指示発話の質的側面の分析は第1試行を取り上げて行った。対象指示コミュニケーション課題において、指示者が行為者との共通基盤をどのように活用しているかを明らかにするため、以下のとおり、南部・原田（1998a）の枠組みを用いて分析を行った。

はじめに、すべての指示発話を命題に基づく“指示ユニット”に分割した。指示ユニットとは、単独でメッセージとしての意味をなす指示発話の最小単位である。例えば、“じゃあね、手ひろげて、なんかこう、正座して跳んでるような感じ”という発話は、“手ひろげて”、“正座して”、“跳んでるような感じ”の三つの指示ユニットに区別された。さらに、すべての指示ユニットをその内容により次の五つのカテゴリー、“図形的描写”、“部分的メタファ”、“全体的メタファ”、“ラベル”、“身体的指示をともなう指示代名詞”

に分類した（Table 1）。なお、“身体的指示をともなう指示代名詞”は本研究において新たに追加したカテゴリーである。これらは、Fussell & Krauss（1989a, 1989b）と同様に、指示発話がどのような共通基盤にも基づくかを反映するものとして、また対話者間の実際の共通基盤に応じて課題達成の効率や効果に変化するものとして設定された。例えば、サッカーに関心が高い指示者が“バイシクルシュートしてるやつ”という指示発話（全体的メタファ）を示した場合、その指示者は行為者とサッカーに関する知識を共有していると推測し、その情報を活用したと考えられる。このような指示は、実際にサッカーの知識を有している行為者にとっては一言で正しく理解できる効率的かつ効果的な指示となりうるが、そうではない行為者にとっては逆に理解しにくいものとなる。一方、図形としての特徴を記述する指示発話（図形的描写）を示した場合には、誰にでも理解されやすい反面、発話量が多く、冗長になるという性質がある。

各カテゴリーの生起頻度について χ^2 検定を行ったところ、対話条件間の偏りが有意であった（ $\chi^2_{(12)} = 48.60, p < .01$ ）。残差分析の結果（Table 2）、ついたて+モニタ条件では部分的メタファが多く（44.9%）図形的描写が少なかった（16.9%）。低雑音テレビ電話条件では図形的描写が多く（29.1%）全体的メタファと指示代名詞が少なかった（19.8%, 1.8%）。高雑音テレビ電話条件では全体的メタファと指示代名詞が多く（30.9%, 5.4%）部分的メタファが少なかった（34.7%）。これらに対し、テレビ電話条件では各カテゴリーの生起数に特筆すべき偏りが示されなかった。

さらに、映像チャンネルの利用可能性が指示発話の内容に及ぼす影響を検討するため、南部・原田（1998a）の4条件（音声については本研究の4条件と同様で、映像チャンネルなしの設定；同室でのついたて越し、別室での電話、電話の音声に高/低の雑音を加えたもの）と、映像チャンネルの有無による比較を

Figure 4. 指示者の発話における平均文節数。

Table 2
第1試行の指示発話における各カテゴリーの生起頻度と残差

	ラベル	全体的メタファ	部分的メタファ	図形的描写	指示代名詞
ついたて+モニタ	23 (.25)	97 (.40)	157 (2.43)**	59 (-3.62)***	14 (0.71)
テレビ電話	29 (-.23)	135 (.80)	175 (-1.32)	129 (1.57)	10 (-1.81)*
低雑音テレビ電話	30 (.23)	91 (-3.97)***	198 (1.93)*	134 (2.73)***	8 (-2.26)**
高雑音テレビ電話	35 (-.21)	177 (2.62)***	199 (-2.61)***	131 (-0.98)	31 (3.23)***

* $p < .10$, ** $p < 0.5$, *** $p < .01$.

個別に行った¹¹。まずついたて+モニタ条件では、図形的描写が少なく全体的・部分的メタファが多かった($\chi^2_{(3)}=28.31$, $p < .01$)。低雑音テレビ電話条件では、図形的描写が多く全体的メタファが少なかった($\chi^2_{(3)}=13.12$, $p < .01$)。高雑音テレビ電話条件では、全体的メタファが多く部分的メタファが少なかった($\chi^2_{(3)}=11.75$, $p < .01$)。一方、テレビ電話条件では映像チャンネルの有無による有意な偏りは示されなかった。これらの結果は、本研究の4条件による χ^2 検定の結果と類似するものである。したがって、本研究の各対話条件における指示発話の特徴は、映像チャンネルの付加と密接に関係することが示唆された。

考 察

ビデオ対話における身体-視覚的相互作用

これまでの研究では、ビデオ対話システムの映像チャンネルは、対面対話と同様に、表情や動作などの非言語情報を伝達する役割を果たすと考えられてきた(レビューとして Whittaker, 1995)。しかしながら実際には、本研究の通常のテレビ電話を用いた対話、ならびに同室のついたて越しでモニタが設置された対話では、映像チャンネルを介して相手を見る行動も、自分の身体を見せる行動もほとんど生起していなかった。その理由として、一つに、本研究が用いた対象指示コミュニケーション課題が言語的相互作用のみにより達成可能な性質であったことが考えられる。さらにもう一つの理由として、ビデオ対話における身体-視覚的相互作用が、対話者自身にとって、対面対話の場合とは異なる比較的困難な認知的課題となっていた可能性が挙げられる。つまり、あえてこのような認知的課題を達成するまでもない状況であったために、音声情報を中心とした対話が行われたと考えることができるだろう。

一方、音声に障害のあるテレビ電話を用いた場合には、映像チャンネルを介した身体-視覚的相互作用がより多く行われていた。これについては、対話者ペア

が障害そのものを共有していたことが重要な意味をもつと考えられる。すなわちペアは、障害を乗り越えて対話を成立させ共有理解を達成するためには音声以外の対話資源を活用する必要があるという認識を共有しており、それゆえ、障害の程度に応じた映像チャンネルの利用が生起したと推察できる。特に高雑音状況下では、“モニタ見てる?”や“こっち見て”など、身体的指示への注意を直接的に喚起する発話も観察されており、映像チャンネルを介した身体-視覚的相互作用は、対話者ペアが意図的、協働的に取り組むべき認知的課題となっていることが示唆されたと言えよう。したがって、仮に障害が一方の対話者のみに生じている場合には、このような映像チャンネルの利用行動は生起しない可能性があるだろう。

さらに、映像チャンネルを介して相手を見る行動の生起は対話条件により変化していたが、いずれの対話条件でも指示者・行為者の役割による差異が生じていなかった点は興味深い。このことは、身体-視覚情報を対話の中に持ち込む、または持ち込まないという選択が、個々の対話者の判断により個別に行われていたのではなく、対象指示における役割を越えて、対話者ペア相互の了解に基づき協働的に達成されていたことを意味すると考えられる。例えば、映像チャンネルがほとんど利用されなかったついたて+モニタ、テレビ電話の両条件においても、課題が開始される直前にはモニタに映る互いの姿を眺めたり手を振ったりする様子が見られたが、課題が開始されると同時に、両者とも直ちにモニタから視線をはずし音声による対話を行っていた。このことから、ビデオ対話において身体-視覚的相互作用を行うことは認知的課題となっていること、対話者はこれを状況に応じながら意図的、協働的に達成していることが明らかである。

ビデオ対話における言語的相互作用

指示者の指示発話については、特に質的側面に対話条件ごとの特徴が示された。まず通常のテレビ電話を用いた条件では、表現に特筆すべき偏りが示されず、またモニタなしの場合ともほぼ同じ内容であった。この理由として、一つに、電話対話についてはその広範かつ長期にわたる使用によってすでに特定の技能や社

¹¹ “身体的指示をとまなう指示代名詞”を除く4カテゴリーを比較した。

会・文化的ルールが確立されており、“対話の場”（原田，1997）の共有が容易であることが挙げられる。対話の場とは、対話者が認識する心理的な対話空間であり、音声のみによって成立する対話の場は、そこに身体-視覚情報を持ち込むビデオ対話に比べ安定していると考えられる。これに加えもう一つの理由として、すでに述べたように、映像チャンネルを介して身体-視覚的相互作用を行うことは対話者にとって認知的負荷の高い課題であることが挙げられる。これらにより、映像チャンネルが利用可能な状況においても電話型の対話が行われやすいと考えられる。

これに対し、同室のついたて越しでモニタが利用可能な条件では、指示対象の構成要素を比喩的に表現する“部分的メタファ”が多く用いられ、このような特徴は映像チャンネルの付加によりもたらされたものであった。これは、同じく映像チャンネルがほとんど利用されなかったテレビ電話条件とは対照的である。このことから、対話相手と物理的空間を共有しながらもついたてにより隔てられ、さらに相手の姿は映像チャンネルにより提示されるという状況は、対話者に、同室での音声のみの対話とも、電話による遠隔対話とも、おそらくは対面による対話とも全く異なる新奇な対話事態としての認識を生じさせた可能性が示唆されたと言える。

さらに、音声に障害のあるテレビ電話においては、雑音の程度により発話に異なる特徴が示された。まず高雑音テレビ電話条件では、“身体的指示をとまなう指示代名詞”に加え、指示対象を比喩により簡潔に表現する“全体的メタファ”が多く用いられていた。これは、身体-視覚情報を対話資源として積極的に利用すると同時に、高雑音により指示発話の細部が聞き逃されるリスクを少なくし、行為者との固有の共通基盤を活用するという対話方略が用いられていたことを示している。

一方、低雑音テレビ電話条件では指示対象を記述的に表現する“図形的描写”が多用されており、低雑音による障害は映像チャンネルの利用を促すと同時に、一般的他者との共通基盤に基づく冗長だが確実な指示を導いていた。このことから、映像チャンネルの利用可能性が認識・共有されている状況においても、実際に身体-視覚情報を対話の中に持ち込むことは容易ではなく、その具体的な方略が模索されていたと考えられる。

また、指示発話の量的側面については、いずれの対話条件においても特に前半の試行で試行ごとに減少する傾向が示された。これは、同じ指示対象について指示・同定を繰り返すことによって対話者間に新たな共通基盤が構築され、共有理解がより効率的に達成可能になることを示した Clark & Wilkes-Gibbs (1986) と一致する傾向である。また映像チャンネルの利用行動

についても、これと同様の減少傾向が示された。したがって、本研究のいずれの対話条件においても、対話者間に新たな共通基盤が蓄積され、これによって課題達成の効率化が促されたと言える。

まとめと今後の課題

以上の考察から、ビデオ対話は、各チャンネルのもたらす情報の単純な加算によっては説明することのできない、対面対話とは本質的に異なる活動として成立していることが明らかになった。これを支えているのは、対話システムの物理的特性に応じた対話者のさまざまな資源の活用や、単なる必要量の情報交換にはとどまらない対話者間の協働的な相互作用であると言える。

本研究では、個々の対話者における映像チャンネルの利用行動と指示者の発話を分析対象として取り上げたが、対話者間の協働の様態をより詳細に検討するために、身体的指示が実際にはどの程度相手に見られていたのか、また行為者はどのような形で対話に関与していたのかなどの分析が今後の課題として残された。さらに、音声の障害が一方の対話者のみに加えられた場合についても検討する必要があるだろう。

ビデオ対話システムをさまざまな領域で有効に活用していくためには、今後さらに、ビデオ対話とはいかなる認知的、協働的過程かをより詳細に検討し、それらを支援するためにシステムのデザインはどうあるべきか、その使用者にとっての“使いやすさ”の実現には何が必要かを考えていくことが重要である。

引用文献

- Clark, H. H., & Wilkes-Gibbs, D. 1986 Referring as a collaborative process. *Cognition*, **22**, 1-39.
- Farr, R. 1991 Bodies and voices in dialogue. In I. Markova & K. Foppa (Eds.), *Asymmetries in dialogue*. Hemel Hempstead: Harvester Wheatsheaf. Pp. 241-258.
- Finn, K. E. 1997 Introduction: An overview of video-mediated communication. In K. E. Finn, A. J. Sellen & S. B. Wilbur (Eds.), *Video-Mediated Communication*. Hillsdale, NJ: Lawrence Erlbaum Associates. Pp. 3-21.
- Fussell, S. R., & Krauss, R. M. 1989a The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Memory and Language*, **26**, 209-225.
- Fussell, S. R., & Krauss, R. M. 1989b Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, **19**, 509-525.
- 原田悦子 1997 人の視点から見た人工物研究 日本

- 認知科学会(編) 認知科学モノグラフ vol. 6 共立出版
(Harada, E. T. 1997 *Cognitive engineering studies* Tokyo: Kyoritsu Shuppan.)
- Heath, C., & Luff, P. 1991 Disembodied conduct: Communication through video in a multimedia office environment. *Proceedings of CHI '91 Human Factors in Computing Systems*, 99-103.
- 柿井俊昭 1997 双方向型 TV を用いたマルチメディア・カウンセリングの基礎的研究 心理学研究, **68**, 9-16.
(Kakii, T. 1997 Characteristics of multimedia counseling: A study of an interactive TV system. *The Japanese Journal of Psychology*, **68**, 9-16.)
- 川浦康至(編著) 1993 メディアコミュニケーション 現代のエスプリ **306**, 9-19.
(Kawaura, Y. 1993 Mediated-communication. *L'esprit D'aujourd'hui*, **306**, 9-19.)
- 南部美砂子・原田悦子 1998a 認知的人工物と対話——対象指示コミュニケーション課題による検討—— 認知科学, **5**, 39-50.
(Nambu, M., & Harada, E. T. 1998a Cognitive artifacts and conversation in referential communication tasks. *Cognitive studies*, **5**, 39-50.)
- 南部美砂子・原田悦子 1998b 認知的人工物が介在する対話——身体-視覚情報の利用をめぐる—— 日本認知科学会第 15 回大会発表論文集, 98-99.
(Nambu, M., & Harada, E. T. 1998b Cognitive artifacts and conversation: Analyzing the use of gestural-visual information. *Proceedings of the 15th Annual Meeting of the Japanese Cognitive Science Society*, 98-99.)
- 南部美砂子・原田悦子 1998c 認知的人工物が介在する対話——主観的評価の検討—— 日本教育心理学会第 40 回総会発表論文集, 312.
- Nambu, M., & Harada, E. T. 1999 Video-mediated referential communication: The relation between visual and linguistic information in telephone conversation with monitor links. *Proceedings of the 2nd International Conference on Cognitive Science and the 16th Annual Meeting of the Japanese Cognitive Science Society Joint Conference*, 681-684.
- 南部美砂子・原田悦子・海保博之 2000 “人工物に支援された対話” 研究の動向と展開——パーソナルビューからみた協働過程の検討へ—— 筑波大学心理学研究, **22**, 71-84.
(Nambu, M., Harada, E. T., & Kaiho, H. 2000 The psychology of mediated interpersonal communication: Shifting the focus from technology to participant collaboration. *Tsukuba Psychological Research*, **22**, 71-84.)
- Whittaker, S. 1995 Rethinking video as a technology for interpersonal communications: Theory and design implications. *International Journal of Human-Computer Studies*, **42**, 501-529.

——2000. 2. 4 受稿, 2001. 12. 1 受理——