

## 遠隔サポートにおける映像チャネル導入の効果

米村俊一<sup>†</sup> 宮本勝<sup>†</sup> 中谷桃子<sup>‡</sup>

**あらまし** 高速ネットワークを用いたIP電話や携帯電話が開発され、これらを用いたビデオ通信サービスが普及しつつある。このような映像通信サービスを利用することにより、テクニカルサポートの現場に映像チャネルを容易に導入できる可能性が広がってきた。本研究では、テクニカルサポートセンタのオペレータが技術的スキルの低いユーザに対して機器のセットアップやトラブル解決などのインストラクションを遠隔で実施することを想定し、そこでのサポートに映像チャネルを導入することがユーザとオペレータとの協同作業の効率、およびコミュニケーション品質にどのような影響を与えるかを実験的に検証した。実験の結果、映像チャネルを導入した場合、作業遂行時間は、音声チャネルのみを用いた場合よりも映像チャネルを導入した場合の方が作業時間が短いという結果が得られた。また、映像チャネルを導入した場合、音声チャネルのみによるコミュニケーションよりもユーザの発話数が少ないという結果が得られた。

### The effect of visual information in collaborative troubleshooting

Shunichi Yonemura<sup>†</sup> Masaru Miyamoto<sup>‡</sup> Momoko Nakatani<sup>‡</sup>

**Abstract** Various communications services are emerging with the spread of the broadband network. This is complicating system operation, and call centers are handling a lot more inquiries. Most calls are made by users with little technical skill and the questions are raised often. Unfortunately, such users take a long time to solve even simple problems. A key problem is that the operator has great difficulty in accurately grasping the trouble event and the operating environment of the system. Moreover, it is not easy for novice users to execute the instructions given by the operator. The increased popularity of IP telephones and vision-capable cellular phones suggests that audio and video channels may be used at the same time for problem solving. We examine the effect of visual information on the effectiveness of the collaboration of a novice user and an operator. We observed an audio group and a video group performing the same task and assessed their performances in terms of execution time and the number of utterances.

#### 1. はじめに

ブロードバンドネットワークの普及に伴って、様々なコミュニケーションサービスが展開されている。特に、電子メールやチャット、Web閲覧、さらにはビデオ通信サービスなど、コンピュータシステムを介したコミュニケーションの利用が大幅に増加しつつある。このような状況において、各種サービスの利用条件や取り扱い方法なども複雑になり、テクニカルサポートセンタへの相談件数が大幅に増加している。

テクニカルサポートでの電話による相談では、技術的スキルの低いユーザが何度も同じような質問をしてくる傾向が見られる。また、そのようなユーザは、1回あたりのサポートにおいて問題解決までに長い時間を必要とし、この間、オペレータはそのユーザ対応で拘

束されるためサポートセンタでの対応効率が低下する。特に技術的スキルの低いユーザにとって、トラブルの状況やシステムの動作環境を言葉で正確にオペレータに伝えることは難しく、また、オペレータから指示された対処操作を的確に実行することも容易ではないのが現状である。

一方、高速ネットワークを用いたIP電話や携帯電話が開発され、これらを用いた映像通信サービスが普及しつつある。このような映像通信サービスを利用することにより、テクニカルサポートの現場に映像チャネルを容易に導入できる可能性が高まってきた。

従来より、コミュニケーションにおける視覚情報の役割を明らかにするため、face to faceの対面コミュニケーションと、例えば電話を用いた音声のみのコミュニケーションを比較する等の研究が行われており、視覚的な環境では会話の途切れを気にしない自由な発話や割り込みが可能であり、また、話し方や発話内容も音声のみの場合とは異なり、円滑なコミュニケーションが可能であることが報告[1][2]されている。しかし、

<sup>†</sup>NTT サイバソリューション研究所  
〒239-0847 神奈川県横須賀市光の丘1-1  
<sup>‡</sup>NTT Cyber Solutions Laboratories  
1-1 Hikarinooka Yokosuka-Shi Kanagawa 239-0847 Japan

例えばテレビ会議システムのようにカメラで捉えた視覚情報を共有するような非対面でのビデオコミュニケーションでは、視覚情報が特定範囲に限定されたり、あるいは対話空間の不連続性等からくる違和感が支配的となるため、視覚情報共有の効果が見られないばかりでなく、映像通信独特の違和感が円滑なコミュニケーションを阻害するといった報告[3][4][5]もある。その一方、例えば作業者が指導者からの遠隔支援を受けながら複雑な三次元物体を組み立てるようなタスクでは、映像通信を導入することによって、作業効率および品質の向上は見られなかったものの、視覚情報の共有によってコミュニケーションがより円滑に進められるという報告[6][7][8][9][10]がある。

本研究では、テクニカルサポートセンタのオペレータが技術的スキルの低いユーザに対して機器のセットアップやトラブル解決などのインストラクションを遠隔で実施することを想定し、そこでのサポートに映像チャンネルを導入することがユーザとオペレータとの協同作業における作業効率、およびコミュニケーション品質にどのような影響を与えるかを実験的に検証した。実験の結果、映像チャンネルを導入した場合、作業遂行時間は、音声チャンネルのみを用いた場合よりも映像チャンネルを導入した場合の方が短いという結果が得られた。また、映像チャンネルを導入した場合、音声チャンネルのみによるコミュニケーションよりもユーザの発話数が少なく、円滑にコミュニケーションが行われているという結果が得られた。

## 2. コミュニケーション実験：被験者とオペレータとの遠隔共同作業による問題解決

遠隔テクニカルサポートへの映像チャンネル導入の効果を明らかにするため、低スキルユーザが遠隔でオペレータからの指示を受けながら機器のセットアップ、およびトラブル解決を行うコミュニケーション実験を実施した。

### 2.1 実験方法

技術的スキルの低い被験者に、PC上で動作するTV電話のセットアップおよびトラブル解決を行わせた。被験者は、家庭におけるPC設置環境を想定した実験室(個室)の机上でTV電話システムのセットアップ、およびトラブル解決を行った。作業を進める際、被験者は別室に設けられた仮想的なテクニカルサポートセンタにアクセスし、オペレータからの遠隔指示を受けながら作業を実施した。被験者が案内された個室の机には、インターネット接続用ルータに接続されたノート型PC、Webカメラ等が置いてあり、被験者はWebカメラをPCに接続しアプリケーションソフトを立ち上げて、TV電話システムを使える状態にしなければならない。ただし、実験で使用するPCには音量調整のミュート機能が働いていて音が出ないというトラブルが仕込まれている。被験者は、セットアップ作業の途中でこのトラブルを解決しながらセットアップを完了しなければならない。被験者には、実験で用いるTV電話の操作方法に関する事前知識、および、音が出ないトラブルの存在に関する告知は全く与えなかった。

### 2.2 被験者

低スキルユーザを想定した被験者として、40歳～55歳までの女性10名を用いた。被験者のスキルレベルは、Webのブラウズや電子メールのやりとりを時々行う程度で、ネットワーク設定やプログラム開発経験などは無い。被験者は、オペレータとのコミュニケーションにおいて音声チャンネルのみを用いる音声群と、音声チャンネルに加えて映像チャンネルも併用した映像群とに5名ずつ無作為に分けられた。一方、テクニカルサポートセンタ(仮想)のオペレータ役として、実際のサポートセンタでの業務経験を有する30歳代の女性1名を用いた。

### 2.3 実験装置

実験では、インターネットにルータを介して接続されたノート型PC、およびUSB接続型のWebカメラを用いた。また、TV電話用アプリケーションとしてドットフォンパーソナルV(NTTコミュニケーションズ)を使用した。机上のノート型PCには、事前にドットフォンパーソナルVがインストールされている。また、TV電話で使用するWebカメラはPCから外された状態になっており、被験者はこのWebカメラをPCに接続して、TV電話のアプリケーションソフトを起動し、システムが正常に動作していることを確認するための通話試験を行わなければならない。テクニカルサポートセンタとの連絡ツールとして、音声チャンネルは内線電話を、映像チャンネルは家庭用のビデオカメラ(NTSC)に9インチCRTモニタを接続したシステム2セットを双方向で設置した。この実験システム構成により、音声チャンネルおよび映像チャンネル共に、遅延がほとんど無いコミュニケーション環境を実現した。

実験者は、被験者の隣の部屋からハーフミラーを介して実験の様子を観察し、被験者に対して必要な指示を与えた。また、実験の様子はデジタルビデオ(MPEG2)で撮影し、実験後の解析で使用した。

### 2.4 実験手続き

被験者は実験用の個室に案内され、最初に実験の概要、および、実験で使用する機器の概要に関する説明を受けた。実験開始後、被験者はテクニカルサポートセンタにアクセスし、それぞれ許されたチャンネルを通してオペレータの指示を受けながらTV電話のセットアップ、およびセットアップのプロセスで遭遇するトラブルの解決を行った。

実験では、被験者とオペレータとの共同作業プロセスを次に示す8つの段階に分けた。

#### (1) PCへのWebカメラの接続：

はじめに、被験者がオペレータにアクセスし、オペレータの指示に従ってWebカメラをPCに接続する。この段階では、被験者は机上にある実験用PCの状況(動作状態や接続機器など)等をオペレータに報告しなければならない。

#### (2) TV電話の起動：

オペレータの指示に従って、ドットフォンパーソナルVを起動(Windowsのスタートメニューから)する。

#### (3) TV電話へのログイン：

平均作業時間

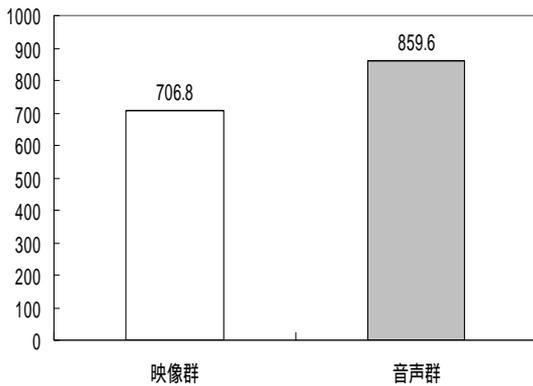


図1 平均作業遂行時間

お客様情報（予め印刷物を机上に配布）を参照しながらキーボードよりパーソナルIDを入力し、TV電話システムにログインする。

(4)ログインの確認：

正常なログインを確認すると共に、TV電話のメイン画面、および自分の顔映像がPCに表示されていることを確認する。

(5)サウンド品質の設定：

TV電話のセットアップに必要な「サウンド設定 & 通話品質確認」を実施する。被験者はこの段階で、オペレータの指示どおり操作しても「音が出ない」というトラブルに遭遇する。

(6)音量調整：

被験者は、オペレータの遠隔指示に従って「音が出ない」トラブルの原因を調査し、解決のための操作を実施する。

(7)通話品質の確認：

トラブル対処策を実施後、再度「サウンド設定 & 通話品質確認」を実施してトラブルが解決したことを確認する。

(8)システムの動作確認：

TV電話が正常に動作することを確認するため、仮想的な友達（実験者）にTV電話をかけて映像通信を行う。

3. 実験結果

音声群と映像群のパフォーマンスの違いは、トラブル解決までの作業遂行時間、および発話数において観測された。

3.1 作業遂行時間

図1は、Webカメラの接続からシステムの動作確認までの8段階の作業を終えるまでの平均作業遂行時間である。縦軸は作業遂行時間（秒）、横軸は2つの実験群（映像群および音声群）である。図1に示すように、8段階の全作業に渡る平均作業遂行時間は、映像群で706.8秒、音声群で859.6秒であり、映像を用いた被験者の方が音声のみを用いた被験者よりも効率的に作業が遂行できた。ただし、8段階の作業全体での2つの平均作業遂行時間の差を分散分析( $F(1,8)=1.485$ ,  $p=0.258>0.05$ )すると、有意な差とはならなかった。

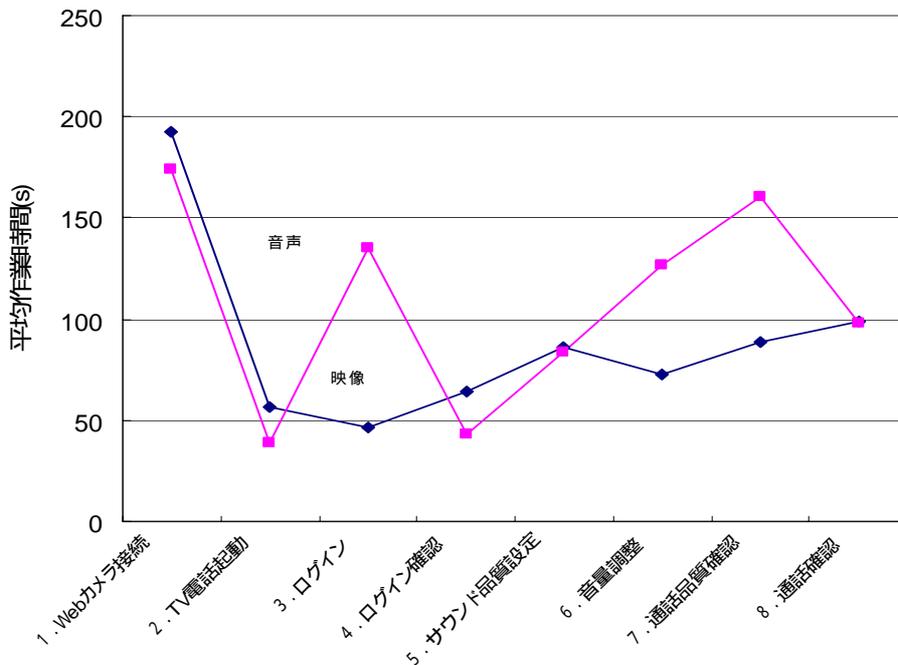


図2 各作業ステップごとの平均作業遂行時間

図2は、図1に示した平均作業遂行時間を各作業ステップごとに示したグラフである。縦軸は作業遂行時間(秒)、横軸は各作業ステップであり、左から右へと作業ステップが進む。図2で、 は映像群のプロットを、 は音声群のプロットを示す。図2に示すように、映像群/音声群の平均作業遂行時間の差が大きいのは3・ログイン、6・音量調整、および7・通話品質確認の各段階であった。分散分析の結果、2つの実験群と作業ステップとの間に有意な交互作用 ( $F(7, 56)=2.45, p<0.05$ )が観測された。

このうち、3・ログインでの差は、長時間を要した被験者のデータに平均値が引きずられたものであるが、トラブル原因を特定し、それに対処するために被験者とオペレータの間で頻繁なやり取りが発生した段階である、6・音量調整、および、7・通話品質確認では映像群と音声群で有意な違いが見られた。以上のことから、特にトラブルの原因を特定したり、その解決を図るような場面においては映像チャンネルの導入が遠隔共同作業の効率向上に寄与することがわかった。

### 3.2 被験者の発話数

図3は、全作業を遂行する間に被験者が発話した単語の総数の平均である。図3において、縦軸は平均発話数、横軸は各実験群(映像群および音声群)である。図3に示すように、実験中の平均発話数は、映像群が113.0語、音声群が187.2語であった。映像チャンネルを与えられた被験者に比べ、音声のみによるコミュニケーションを行った被験者は1.7倍も多くの発話を行ったことになる。分散分析の結果、映像群と音声群での平均発話数の差( $F(1, 8)=11.28, p<0.01$ )は有意であった。すなわち、音声チャンネルのみでの遠隔協同作業は、映像チャンネルを利用する場合にくらべて多くの発話を行っていることがわかる。実際、実験中の被験者を観

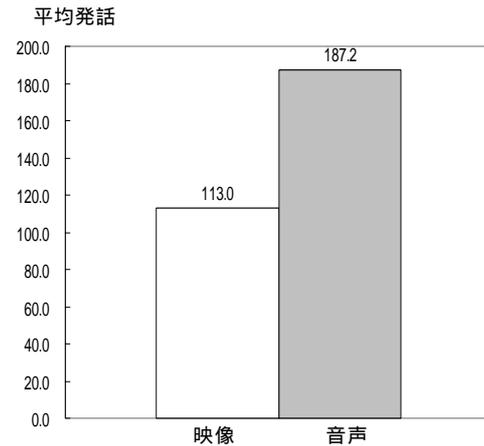


図3 平均発話数

察していると、映像チャンネルを使用する被験者は自らのPCの状況をオペレータに伝えるのにビデオ情報の存在を前提とした指示語を用いていた。例えば、Windowsの音量調整アイコンの場所を探す際、映像群ではオペレータのPC画面を遠隔映像で確認しながら「ん、どこですか? ああ、はいはい。」という発話のに対し、音声群では「スピーカー? スピーカですか? ... いえ、まず、あの、時刻が出てますよね...」といったように電話を用いて多数の発話を行った。

図4は、各作業段階における平均発話数である。図4において、縦軸は被験者が発話した平均単語数を、横軸は各作業の段階を示す。図4で、 は映像群のプロットを、 は音声群のプロットを示している。図4に示すように、ステージ2およびステージ4を除き、音声群の発話数は映像群の発話数を上回っているのがわかる。分散分析の結果、各実験群と作業段階の間に交互作用( $F(7, 56)=24.027, p<0.01$ )が観測された。特に、

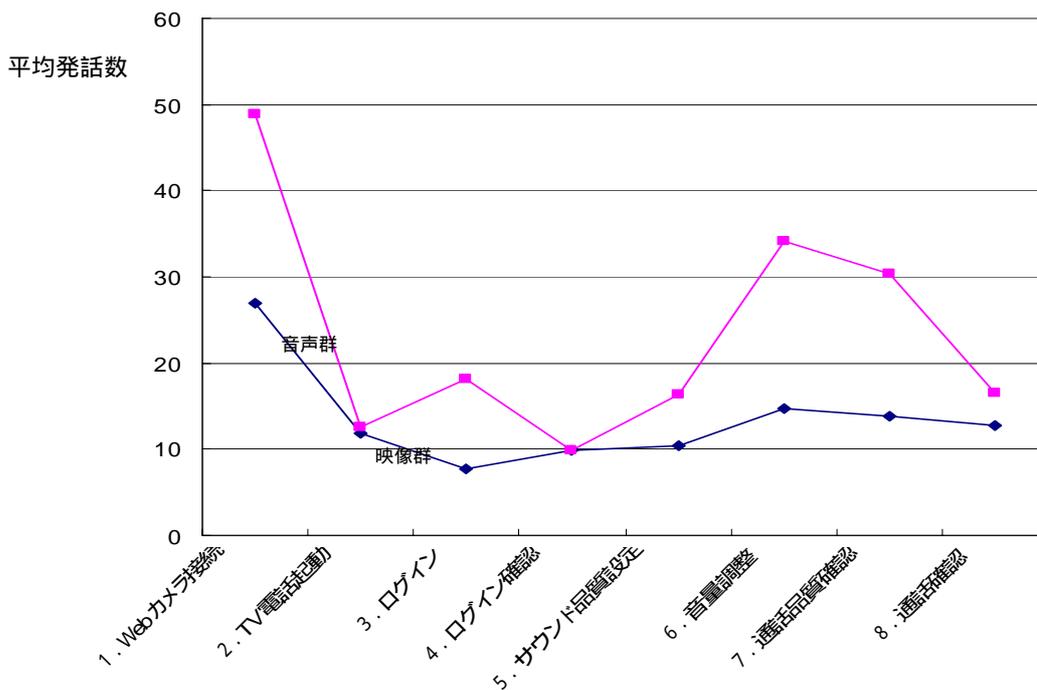


図4 各作業段階における平均発話数

ステージ1、ステージ6、およびステージ7において発話数の大きな差が見られた。ステージ1は、被験者とオペレータのコミュニケーションが始まる段階であり、ここでは被験者は実験用PCの動作状況や周辺機器に関する情報をオペレータに伝達をしなければならない。また、前述のとおり、ステージ6、およびステージ7では被験者のPCで発生したトラブルの原因をいち早く特定し、それに対処するために被験者とオペレータの間で頻繁なやり取りが発生した段階である。発話数の大きな差は、まさにこの段階で発生した。特に音声チャンネルのみを用いた被験者の場合、オペレータの発話内容を確認するため、オペレータの発話内容を復唱したり、あるいは、何度も聞き返すといった行動が見られた。

#### 4. 考察

機器のセットアップやトラブル解決を遠隔共同作業によって効率的に実行するためには、ユーザ側の機器の環境やそこで発生しているトラブル状況などをオペレータが正確に把握し、適切な対処の手順を特定すると共に、ユーザがオペレータの指示どおりに対処手続きを実行することが必要である。しかし、特にユーザのスキルレベルが低い場合、機器で発生している状況を明確に説明するためのテクニカルな用語を知らなかったり、また、日常使用しているアプリケーションの基本操作以外のシステム操作に関する知識を持っていない場合が多い。

テクニカルサポートによる支援を受けながら遠隔地にいるユーザが効率よくトラブルを解決するためには、

ユーザとオペレータとの間で交わされる次のようなプロセスが円滑に進行することが必要である。

1. ユーザによるトラブル状況報告：ユーザは、ユーザ側システムで発生したトラブルの状況およびシステムの動作環境等に関する情報を正確にオペレータに伝える。
2. オペレータによるトラブル状況把握：オペレータは、ユーザからの報告に基づき、発生したトラブル状況、およびシステムの動作環境を正確に把握する。
3. オペレータによるトラブル対処手続きの特定：オペレータは、ユーザからの状況報告に基づいてトラブル対処の手続きを特定し、それをユーザに対して正確に伝える。
4. ユーザによるトラブル対処の実行および状況報告：ユーザは、オペレータからの指示を正確に実行し、指示を実行した結果システムがどのような状態になったかを正確にオペレータに伝える。

しかし、特に技術的スキルの低いユーザにとって、トラブルの状況やシステムの動作環境について言葉で正確にオペレータに伝えることは難しく、さらに、オペレータから指示された内容を的確に実行することも容易ではない。このため、テクニカルサポートセンタにユーザがアクセスしてきても、(1)ユーザはトラブルの状況をオペレータに正確に伝えられない、あるいは、(2)オペレータからの指示内容の理解が不十分ためオペレータが意図した通りの操作が実行できない、といったようなことが繰り返される。したがって、ユーザの側ではトラブルの状況を報告したりオペレータか

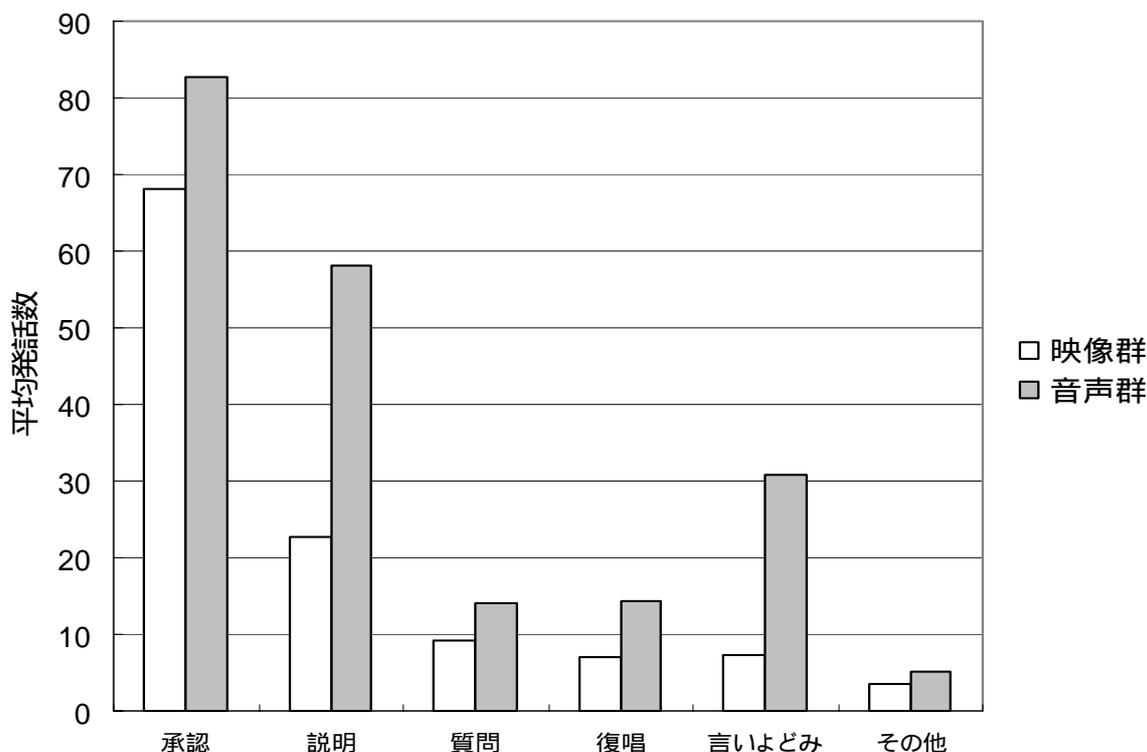


図5 被験者の発話内容の分類

らの指示を理解するのに多大な労力を費やすが、その割には、オペレータ側ではユーザ側で発生している状況の認識が不確実で効果的な対処策が特定できない。このような繰り返しがユーザおよびオペレータ双方のストレスを高めるとともに、トラブル復旧までに多大な時間を費やすという悪循環につながっている。

図5は、被験者の発話内容を「はい、いいえ」といった「承認」、トラブル状況を報告するなどの「説明」、オペレータに対する「質問」、オペレータの発話内容を自ら言いなおす「復唱」、明確な発話とならず「えーと、んーと」などと口ごもるような「言いよどみ」、「その他」に分類し、一人当たりの発話数として平均をとったものである。図5に示からわかるように、特に「説明」と「言いよどみ」において音声群の被験者は映像群よりも多くの発話をしている。これは、音声チャンネルのみでは自らの機器の状態を的確に説明できないため、ひとつの事象について何度も繰り返し、言葉を変えながら説明していることを示している。さらに、「言いよどみ」の多さは、発話の際になかなか適切な言葉が出てこないため、頻繁に言いよどんでしまうことを示している。

これと対照的に、映像チャンネルを用いた被験者の場合、自らの機器状態を映像によって直接オペレータに提示できるため説明的な発話は少なく、かつ、言いよどみも少ない結果となった。実際、映像群の被験者では、オペレータに送る映像を写すカメラを小脇に抱えたまま離さないといった行動が見られた。また、発話内容においても、「これ」「それ」などの直接(deictic)的な指示語を使用する頻度が高かった。このことは、技術的スキルの低いユーザの場合にテクニカルサポートの効率および品質に大きく影響する語彙の無さに起因する問題を映像チャンネルで解決可能であることを示している。

以上のように、遠隔で行うテクニカルサポートにおいて、特にエンドユーザの技術的スキルレベルが低い場合、コミュニケーションチャンネルとして映像を導入することによってオペレータとの円滑なコミュニケーションを行うことができ、その結果、効率的な遠隔共同作業が可能となることが示された。

## 5. まとめ

本論文では、テクニカルサポートセンタのオペレータが技術的スキルの低いユーザに対して機器のセットアップやトラブル解決などのインストラクションを遠隔で実施することを想定し、そこでのサポートに映像チャンネルを導入することがユーザとオペレータとの協同作業における作業効率、およびコミュニケーション品質にどのような影響を与えるかを実験的に検証した。実験の結果、映像チャンネルを導入した場合、作業遂行時間は、音声チャンネルのみを用いた場合よりも映像チャンネルを導入した場合の方が短いという結果が得られた。また、映像チャンネルを導入した場合、音声チャンネルのみによるコミュニケーションよりもユーザの発話数が少なく、円滑なコミュニケーションが行われているという結果が得られた。

今後、ブロードバンドサービスの進展に伴ってテクニカルサポートセンタへの問い合わせが激増することが予想されている。すでに、多くのサポートセンタで

はユーザからの問い合わせに答えきれない状況が発生しており、今後はより一層の対応効率の向上が望まれる。そのためには、本論文で示したような新たなメディアを効果的に用いることによってユーザとのコミュニケーションを適切に行い、質の高いサポートを実現することが必要である。

## 参考文献

- [1] D. R. Rutter and G. M. Stephenson: The role of visual communication in synchronising conversation, *European Journal of Social Psychology*, 7, pp.29-37, 1977.
- [2] D. R. Rutter, G. M. Stephenson and M. E. Dewey: Visual Communication and the content and style of conversation, *British Journal of Social Psychology*, 20, pp.41-52, 1981.
- [3] 原田悦子, 認知工学から見た通信メディア: 対話という課題, 情報処理学会研究報告, 情報メディア 16-2, pp.9-16, 1994.
- [4] 原田悦子, 受話器の心理学的効果: インタフェースとしての受話器, 現代のエスプリ 306, pp.75-83, 1993.
- [5] 原田悦子, 南部美砂子, 認知的課題としての対話: 遠隔対話の心理学的分析から, 人工知能学会第14回全国大会, S2-02, pp.60-61, 2000.
- [6] R. E. Kraut, S. R. Fussell, and J. Siegel: Visual Information as a Conversational Resource in Collaborative Physical Tasks, *Human-Computer Interaction*, vol.18, pp.13-49, 2003.
- [7] S. R. Fussell, L. D. Setlock and R. E. Kraut: Effects of Head-Mounted and Scene-Oriented Video Systems on Remote Collaboration on Physical Tasks, *CHI2003*, vol.5, No.1, pp.513-520, 2003.
- [8] E. S. Veinott, J. Olson, G. M. Olson and X. Fu: Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other, *CHI99*, pp.302-309, 1999.
- [9] O. D. Jones, A. Monk and L. Watts: Some advantages of video conferencing over high-quality audio conferencing: fluency and awareness of attentional focus, *International Journal of Human-Computer Studies*, vol.49, pp.21-58, 1998.
- [10] A. J. Sellen: Speech Patterns in Video-Mediated Conversations, *CHI92*, pp.49-59, 1992.