

研究論文

聴覚障がい者とその家族のためのテレビ視聴時 コミュニケーション支援システムの提案

平尾 美佐^{1,a)} 石井 陽子^{1,3} 宮崎 泰彦¹ 松畷 信貴² 小林 透^{1,4}

受付日 2012年12月14日, 採録日 2013年4月26日

概要: 地上デジタル放送やIPTVなど映像サービスに関する技術革新にともない、映像視聴のスタイルが多様化している。その一方で、聴覚障がい者の映像アクセシビリティへの対応が課題となっている。本研究では、聴覚障がい者を含む家族のテレビ視聴時のコミュニケーションを支援するシステムの構築を目指している。インタビュー調査の結果から、聴覚障がい者は得る情報が視覚情報に限られるため、テレビ視聴において聴覚障がい者特有の制約があることが分かってきた。そこで本稿では、テレビを見ているユーザの手話や表情の映像をテレビ画面上で番組映像とともに表示する「一緒に見ている人の顔が見えるテレビインタフェース」を提案する。本インタフェースのプロトタイプを用いた映像視聴実験を行ったところ、本インタフェースを使用しないときと比べて、発話数が増える、発話の長さが長くなるなどの結果が得られた。

キーワード: コミュニケーション, 聴覚障がい者, アクセシビリティ, テレビ, ユーザインタフェース

Proposal of TV Communication Support System for the Hearing Impaired and Their Families

MISA HIRAO^{1,a)} YOKO ISHII^{1,3} YASUHIKO MIYAZAKI¹
NOBUTAKA MATSUSHIMA² TORU KOBAYASHI^{1,4}

Received: December 14, 2012, Accepted: April 26, 2013

Abstract: The TV watching style has been diversified thanks to the innovation of video viewing services such as the integrated services digital broadcasting and IPTV. On the other hand, the TV accessibility of hearing-impaired people has been still problems. We aim to construct the communication support system for the family who has hearing-impaired people in time of watching the TV. Hearing-impaired people have some unique limitations in case of watching TV because hearing-impaired people rely on mostly the visual information. Therefore, we propose “TV interface which combines TV program and watching people’s face”. This new interface allows us to see TV program and watching people’s sign language or expression simultaneously. We show the TV watching experiment results using our proposed interface prototype system. These results indicate that the communication frequency has been increased rather than the case of not using this prototype system.

Keywords: communication, hearing impaired, accessibility, TV, user interface

¹ 日本電信電話株式会社 NTT サービスエボリューション研究所
NTT Service Evolution Laboratories, NTT Corporation,
Yokosuka, Kanagawa 239-0847, Japan

² 電気通信大学大学院情報理工学研究科情報・通信工学専攻
Department of Communication Engineering and Informatics,
Graduate School of Informatics and Engineering, The
University of Electro-Communications, Chofu, Tokyo 182-
8585, Japan

³ 2013年2月 日本電信電話株式会社辞職

⁴ 2013年3月 日本電信電話株式会社辞職 現 長崎大学工学部工
学科

1. はじめに

近年の技術革新にともない、地上波テレビ放送のデジタル化が進み、インターネットを利用したテレビ向け映像配信サービスも増加している。お気に入りのアーティストのコンサートを高画質・高音質の中継映像で楽しむ、VODを

School of Engineering, Nagasaki University

^{a)} hirao.misa@lab.ntt.co.jp

利用して好きな映画を好きなときに楽しむなど、テレビの楽しみ方は多様化している。

一方で聴覚に障がいがあるユーザにとっては、満足が得られる映像視聴環境が実現されているとはいえない。情報保障の観点での取り組みはこれまでも多く行われており [1], [2], 字幕放送に対応する番組の割合は増える傾向にあるが [3], テレビ視聴時のユーザエクスペリエンスを高めるためには、コンテンツの内容を漏れなく伝え正しい理解を促すだけでは十分ではない。テレビはコンテンツ消費のためのツールという以外にも、家族団らんのためのコミュニケーションツールという側面を持つからである。そこで筆者らは、聴覚障がい者を含む家族のテレビ視聴時のコミュニケーションを支援するシステムの構築を目指している [4]。

テレビ視聴時に行われるコミュニケーションは、家庭によって、また見ている番組によって様々であるといえる。バラエティ番組を見ながら出演しているタレントのゴシップで盛り上がることもあれば、旅番組を見ながら家族で行った旅行の思い出を振り返ることもあるだろう。他に、見ているテレビ番組の内容とはまったく関係のない会話をすることもある。食事をしながらテレビを見ているのであれば食卓に並んだおかずが美味しいと言って料理した人を褒めることもあるだろうし、夫婦が子どもをあやしながらテレビを見ているのであれば、子どもを抱いている妻に対して夫が「今度は僕が抱っこするよ」などと声をかけることもあるだろう。

自身と一緒にいる相手がテレビを見ている状況下でわざわざコミュニケーションをとろうとするのはなぜだろうか。先にあげた例では、見ているテレビ番組に関連したコミュニケーションを行っている場合と、テレビの内容とはまったく関係のないコミュニケーションを行う場合に分けられる。前者の場合は、テレビ番組を見ながら感じたこと、思い出したことなどについて、相手に知ってほしい、共感してほしいという気持ちがあるからではないだろうか。また後者の場合は、そのときどうしても相手に伝えたい・伝えなければならないことがあるからではないだろうか。

前者の場合も後者の場合も、高いユーザエクスペリエンスをもたらすには、相手に知ってほしい、共感してほしい、伝えたい、伝えなければならないと思うことを自由に伝えられる状態となる必要があるといえる。

そこで筆者らは、構築するコミュニケーション支援システムが実現すべき状態を、テレビを視聴しながらも「相手に知ってほしい、共感してほしい、伝えたい、伝えなければならないと思うことを聴覚障がい者特有の制約を受けずに伝えられる状態」とする。なお、「聴覚障がい者特有の制約」については 2.1 節で示す。

本稿ではまず 2 章で、聴覚障がい者を含む家族のテレビ視聴時コミュニケーションの現状をインタビュー調査の結

果を用いて明らかにし、それをもとに提案するシステムの要件について述べる。3 章では関連研究をあげ、本研究との相違点を明らかにする。4 章では、基本機能の検討とそれに基づいたプロトタイプの実装、5 章ではプロトタイプを用いて行った映像視聴実験、6 章では考察を述べ、7 章でまとめる。

2. 要件抽出

2.1 聴覚障がい者を含む家族のテレビ視聴時コミュニケーションの現状

聴覚障がい者を含む家族がテレビを視聴する際、どのようにコミュニケーションを行うのか、また、コミュニケーションを行ううえでどのような問題点があるのかを具体的に把握するため、聴覚障がい者同士の夫婦 6 組、聴覚障がい者と聴者（以下では、聴覚に障がいがない人のことを聴者と呼ぶ）の夫婦 6 組、聴覚障がい者と聴者の親子 3 組、計 15 組にインタビュー調査を行った。

まず、互いにコミュニケーションをとる際の方法についてたずねると、口話または手話を使用している組、もしくは、口話と手話を場面に応じて使い分けたり、組み合わせ使用したりする組とがいた。口話とは、聴覚障がい者が相手の口の動きを見て言葉を読み取ること（読話）および、口の動きで意思を伝えるコミュニケーション手法である。一方の手話は、3次元空間上で身体を使って表される言語で、大きく手指信号 (Manual Signs) と非手指信号 (Non-manual Signals) で構成されている。手指信号は手の形、位置、運動で表示され、主に語を形成する役割を果たし、非手指信号は表情、視線、頭の傾きなどの要素で構成され文法的機能を果たすといわれている [5]。手話観察者の視線計測に関する先行研究 [6] では、手話母語者の視線がほとんど顔に集中することが確認されている。このことから、手話コミュニケーションにおいて顔の表情や頭の動きといった非手指信号の重要性が分かる。

次に、聴覚に障がいがある回答者に、テレビ視聴時に利用している情報保障の手段についてたずねると、全員がテレビのクローズドキャプション機能を利用し、映像と字幕から情報を得ていた。これは、1人でテレビを視聴するときも家族と視聴するときも同様であった。このことから、聴覚障がい者は家族とのコミュニケーションもテレビ視聴も、どちらも視覚情報により行っていることが分かる。

聴者のみより構成される家族が団らんの場でテレビを視聴しているシーンを想像してみると、目ではテレビ番組を見ながら、耳では家族の話聞き会話をするという行為は何気ないものに思える。しかし、家族とのコミュニケーションもテレビ視聴も、どちらも視覚情報により行っている聴覚障がい者やその家族がその両方を並行して行うことは容易ではない。

ただ、なかには長年連れ添っている夫婦など、慣れてい

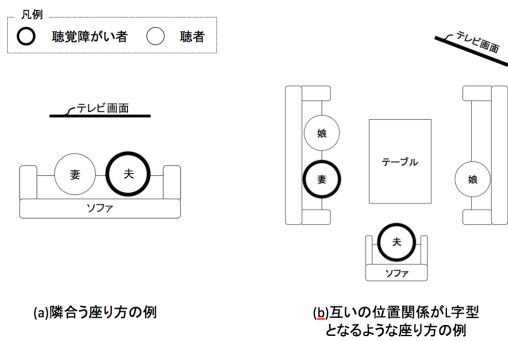


図 1 インタビュー回答者のテレビ視聴時の座り方を再現した模式図
Fig. 1 The diagram shows how the respondents sit when they watch TV.

る相手とのコミュニケーションであれば、番組を見ながら互いにコミュニケーションがとれるという回答者も複数組いた。そのうち 1 組の夫婦の回答内容を以下に示す。

夫 (聴覚障がい者) : 「まったく知らない人と隣り合った場合にはその人の表現がどうなるのかってというのは分からないですね。ただ、妻との関係は長いわけですから、顔を見なくても手が動いたことで相手の気持ちも分かるんです。」

妻 (聴者) : 「手話で話すときに、別に完全にお互いを見なくても、この辺で手が動いてるとだいたい何言ってるかさんな難しいことでなければ分かる。見たがってそうときとか、あとドラマとか筋がすごく大事なときとかはそんな長い会話はしないっていうのはありますけど。」

ここで紹介した夫婦は、必ずしも相手と面と向き合っていないけれども、視界の隅に入ってくるわずかな手の動きを通して会話ができるようである。ただし、これはテレビ視聴の際、互いの手話が少しでも視界に入るような位置関係で座っていることが前提となる。実際に、この夫婦や同様の回答をした回答者たちは、テレビ視聴の際には隣り合って座る、互いの位置関係が L 字型となるように座るなど (図 1)、テレビ画面を見ながらも互いの手話が少しでも見やすくなるよう座る位置の工夫を行っていた。

さらに前述の妻の話からは、手指信号や非手指信号がはっきりとは見えない状態でできるのは簡単な会話に限られることがうかがえる。また、相手のテレビ視聴を邪魔しないようにという気遣いの意味からも会話を短く済ませる様子が読みとれる。

以上のことから、聴覚障がい者を含む家族がテレビ視聴時に次のような制約の中でコミュニケーションを行っていることが分かった。

- ① テレビ画面を見ながらも互いの手話が少しでも見やすくなるよう座る位置の工夫を行う。
- ② 相手のテレビ視聴を妨げないよう会話を短く済ませる。

2.2 一緒に見ている人の顔が見えるテレビインタフェースの要件

以上で述べた聴覚障がい者を含む家族のテレビ視聴時コミュニケーションの現状をふまえ、一緒に見ている人の顔が見えるテレビインタフェースを提案する。聴覚障がい者を含む家族がテレビを視聴しながらも、相手に知ってほしい、共感してほしい、伝えたい、伝えなければならないと思うことを聴覚障がい者特有の制約を受けずに伝えられる状態を実現するため、本インタフェースにより 2.1 節で述べた制約①および②を取り除くことを目指す。そこで、聴覚に障がいがあるユーザのテレビ視聴時の視点を増やすことにより、聴者が目ではテレビ番組を見ながら、耳では家族の話の聞き会話をするのに近い状態を実現する。具体的には、テレビを視聴しているユーザをテレビ画面側からカメラで撮影し、顔の表情や頭の動きと手話の可動域である上半身をとらえる映像 (以下では、この映像をユーザ映像とする) を新たな視点としてテレビ画面内の番組映像に重畳して表示する。

3. 関連研究

3.1 手話映像の評価に関する研究

本研究に関連する研究の 1 つとして、テレビ電話を利用した手話の遠隔通信を想定し、手話映像の品質評価を行っている中園らの行っている研究があげられる。文献 [7] では、手話映像を 2, 4, 12 インチのサイズで表示し、その手話を文字で書き取らせる実験について報告されている。その結果、実験で使用した 2~12 インチの範囲であれば、表示画面のサイズは可読性には影響がないことが確認されている。ただし中園らは、この結果には実験に用いたユーザ映像のフレームレートが 30 fps と十分に高かったことも影響している可能性があるとして自ら指摘している。

中園らの実験は、ノート PC に手話映像を表示して被験者に読み取らせるものだった。本研究において提案する一緒に見ている人の顔が見えるテレビインタフェースはテレビ画面上に頭部や上半身をとらえるユーザ映像を表示するものであり、画面とユーザとの距離は対ノート PC のときと比べて大きく離れることになる。そのため、テレビ画面上での手話の読み取りに適したユーザ映像のサイズは、別途検証する必要がある。

3.2 聴覚障がい者のための教室におけるコミュニケーション支援システム

ユーザ映像を用いたコミュニケーション支援システムの先事例の 1 つとしては、Takeuchi らが提案する教室におけるコミュニケーション支援システムがあげられる [8]。

聴覚障がいがある学生を対象とした授業において、学生が手話で質問をしても、他の学生は質問者の手話が読み取りにくいという問題があった。そこで Takeuchi らのシス

テムは、教室前方に設置したカメラが、学生が挙手したことにより質問者を特定、質問者にズームし、その映像を教室前方のモニタに表示することで、学生らが質問者の手話を読み取ることを可能とした。

この教室におけるコミュニケーション支援システムと一緒に見ている人の顔が見えるテレビインタフェースとは、聴覚に障がいのあるユーザにユーザ映像を提示し、コミュニケーションを行う際の視点を新たに増やすというコンセプトは共通している。しかし、Takeuchiらのシステムは、質問者から他の学生へ向けた1対多かつ、1方向のコミュニケーションを前提としており、この点では、1対1以上の双方向コミュニケーションを前提としている本研究と大きく異なる。さらにTakeuchiらのシステムは、授業に関する質問内容の共有が目的であるため、システム利用時には、情報伝達を素早く、正確に行うことができる状態となることが望ましいといえる。これに対して一緒に見ている人の顔が見えるテレビインタフェースは、テレビ視聴時に会話することを目的としている。そのためには、コミュニケーションの迅速性や正確性もさることながら、1章でも述べたように、相手に知ってほしい、共感してほしい、伝えたい、伝えなければならないと思うことを聴覚障がい者特有の制約を受けずに伝えられる状態の実現が求められる。

4. 基本機能の検討と実装

4.1 基本機能の検討

一緒に見ている人の顔が見えるテレビインタフェースで実装すべき基本機能について検討するため、聴覚障がい者同士のペア2組、聴覚障がい者と聴者のペア3組を対象に予備検討を行った。なお、この5組のペアは職場の同僚同士で、テレビを見ながら気兼ねなく会話ができる間柄であった。予備検討では、ユーザ映像としてテレビ画面の近傍に設置したウェブカメラで対象者を撮影し、その映像とテレビ番組映像とをノートPCで擬似的に重畳、テレビ画面に出力するという簡易的なシステムを使用した。テレビには、37型フルHD液晶テレビを使用した。対象者にはこのシステムを用いてテレビ番組を視聴しながら、ユーザ映像を通して口話または手話による会話を行ってもらった。この予備検討により下記の基本機能を抽出した。

- フレームレート

予備検討の中では、ユーザ映像のフレームレートを数種類試した。すると、10fps程度の低いフレームレートでは、対象者によっては手話の可読性が落ちたり、もしくは、読み取りに問題はないが、相手と手話による会話をする際に間合いがつかみにくいといった違和感を感じる対象者がいた。手話可読性を上げるためには高いフレームレートが必要との知見もあり [7]、本研究では30fpsのフレームレートを確保することとする。

- 映像サイズ

予備検討において、ユーザ映像の表示サイズとテレビ番組映像の表示サイズは対象者の要求に応じてそのつど変更した。ユーザ映像のサイズについては、4:3の横長の映像で、長辺を15cmから35cmの幅で変更を可能とした。サイズの要求には個人差があり、特に手話の熟練度が低い聴者の対象者や聴覚に障がいがある対象者の中でも成人後に難聴となった対象者は相手を映し出す映像のサイズが大きければ大きいほど手話が読み取りやすくなる様子が見られた。一方で、生まれつき聴覚に障がいがある手話の熟練度が高い対象者にとって、ユーザ映像のサイズは手話の読み取りやすさに影響していなかった。

番組映像のサイズについては、テレビ画面いっぱいに広げるよりも、少しサイズを小さくすることで、対象者を映し出す映像との重なりを極力少なくする表示方法を好む対象者が多かった。

以上から、プロトタイプの実装にあたっては、ユーザ映像、番組映像のそれぞれについて、サイズをユーザの好みに応じて自由に調整できる機能を設けることにした。

- ユーザ映像の切り出し

予備検討で、ウェブカメラで撮影した映像全体を手話が読み取りやすいサイズで表示しようとするとう番組映像が小さくなってしまおうという問題が指摘された。そこでカメラ映像の中から手話の読み取りに必要なユーザの映像を切り出すこととした。その際、カメラ映像の中からリアルタイムでユーザの顔画像を検出しユーザ映像を抽出する機能を実現することで、ユーザが自由にテレビの閲覧位置を決められるようにした。

予備検討で、隣りあいながら近接し座っているペアの様子を観察していると、互いに触れながら会話をする様子や、相手の手の動きにもう片方が合わせるような動作がよく見られた。対象者からも、位置が近い人とは実空間での位置関係も把握したいといった意見が寄せられたことから、ある一定の距離以内で隣りあう者同士の映像は、同じ領域で切り出しユーザ映像とすることが望ましいと考える。

また、口話を主としてコミュニケーションを行っていた組の対象者からは、頭部から上半身までをとらえたユーザ映像では口の動きが読みにくいという指摘が寄せられた。そこで、頭部から上半身までをユーザ映像として抽出する手話モードと、顔部分のみをユーザ映像として抽出する口話モードを実装することとした。

- 各映像のレイアウト

予備検討においては、対象者を映す映像とテレビ番組映像のレイアウトも対象者の要求に応じてそのつど変更した。対象者からは、視線の移動をできる限り少なくするため、対象者を映す映像と番組映像とをできる限り近接させて表示させたいという要望が多くあった。対象者を映す映像を表示する具体的な位置としては、番組映像の字幕が表



図 2 各モード時のユーザ映像

Fig. 2 User image in sign mode and oral mode.

示されることが多い映像下部を好む対象者と、番組映像の中でも被写体の顔など主要な部分が映る映像中央部にできる限り近接させて表示させたいという対象者がいた。このことからプロトタイプでは、ユーザ映像と番組映像は必ず同一画面内に表示し、それぞれの位置はユーザの好みに応じて自由に調整できる機能を実装することにした。

● 鏡像表示

予備検討においては、ウェブカメラで対象者を撮影した映像を特別な処理は施さずに表示するパターンと、左右反転させる処理を施し鏡像表示するパターンとを試した。筆者らは予備検討実施前、鏡像表示をすると手話が読み取りにくくなるのではないかと予測していたが、実際に予備検討を実施してみると、鏡像表示を好む対象者が多かった。その理由としては、手話は左右反転しても可読性に影響はなく、手話の中で指差しの動作を行う際（たとえば、番組映像に登場する人物について「あの人」と言って話をするときなど）には、鏡像表示であれば、自分の指差した方向と画面の中に表示されている自分の指差す方向が一致するため違和感がないことがあげられた。このことから、ユーザ映像は鏡像表示する機能が必要であるといえる。

4.2 プロトタイプの実装

予備検討により抽出した基本機能を備えたプロトタイプを、HTML5でのCanvas仕様に基づきjavascriptで実装した。リアルタイムで映像を取得・表示可能なWebRTC APIを用い、PCにローカルで接続されたウェブカメラ(1,920×1,080 pixel, 30 fps)の映像をユーザ映像として表示する。本プロトタイプはブラウザ上での表示を想定しており、PCの映像はHDMI通信によりテレビ画面に表示可能である。

以下にソフトウェアの処理を述べる。まず、ウェブカメラの映像の中からテレビ画面前にいるユーザの顔画像をリアルタイムで検出する。ユーザ映像が口話モードの場合、検出された顔画像をユーザ映像として用いる。ユーザ映像が手話モードの場合は以下の処理を行う。検出された顔画像を中心とした顔画像の2倍の幅をユーザ映像の幅とし、顔画像を一番上とした顔画像の3倍の高さをユーザ映像の高さとし、ユーザ映像を抽出する。それぞれのモードにおけるユーザ映像のキャプチャを図2に示す。

ユーザが複数人いる場合は、抽出されたそれぞれの顔画

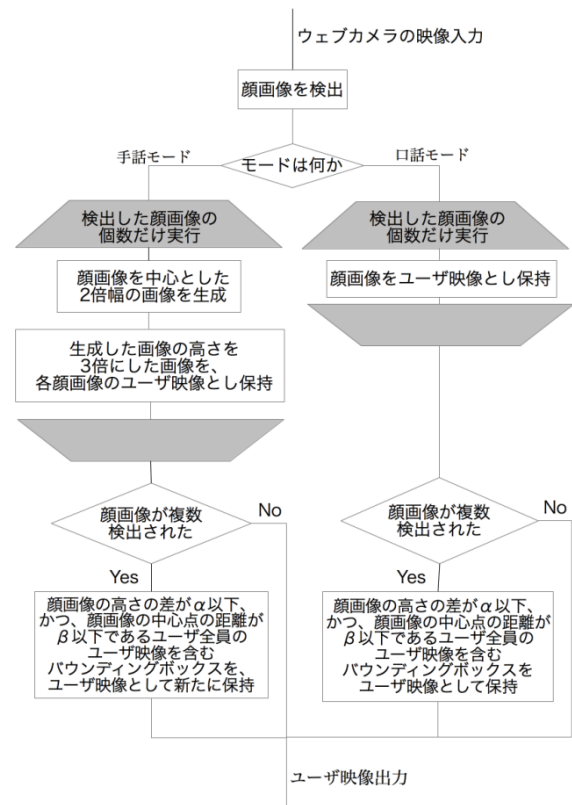


図 3 フローチャート

Fig. 3 Flow chart.



図 4 基本の画面構成

Fig. 4 Basic screen structure.

像の高さの差 α 、および、顔画像の中心点の距離 β を比べる。顔画像の高さの差、および、顔の中心点の距離がそれぞれ設定した閾値以下であった場合、それらユーザ全員のユーザ映像を含むバウンディングボックスを、ユーザ映像として抽出する。本処理の流れを図3に示す。

ユーザ映像は鏡像表示のON/OFFが可能である。ユーザ映像は番組映像とともにサイズ・位置を自由に調整できるよう、jQuery UIによるマウスのドラッグ・ドロップ操作が可能である。基本の画面構成を図4に示す。

さらに、ユーザ映像、番組映像の位置を保存し自由に保存データを読み出せるように図5のようにレイアウト保存画面を設けた。図5では画面が9分割されており、それぞれの領域にユーザが自由に設定した画面構成を可視化し

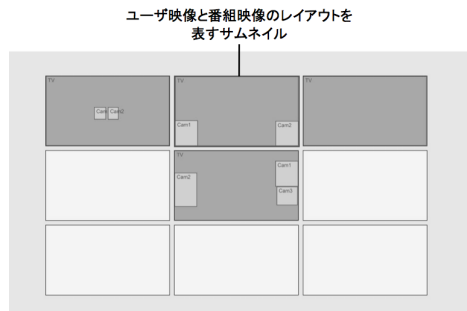


図 5 レイアウト保存画面
Fig. 5 Save layout screen.

ている。画面内をクリックすると、選択された画面構成で本システムを利用可能である。

5. プロトタイプを用いた実験

5.1 目的

プロトタイプにより、2.1 節で述べた制約①および②を取り除くことが可能であるかを検証することを目的として、プロトタイプを用いた映像視聴実験を行った。

5.2 方法と手順

本実験では、聴覚障がい者同士の夫婦 3 組、聴覚障がい者と聴者の夫婦 1 組、聴覚障がい者と聴者の親子 1 組の計 5 組を被験者とした。本研究は口話、手話、および、口話と手話を互いに使う会話でのコミュニケーションを対象としており、これらについて実験を実施し評価する必要がある。しかし、当事者同士しか理解できないことが多い口話によるコミュニケーションの場合、第三者が会話内容を読み取ることが難しい場合がある。本実験では長時間での映像視聴を行うため、映像視聴中の様子をビデオで撮影し後日第三者が詳しく会話内容を読み取る形で記録した。よって、会話内容の正確性を担保するため口話によるコミュニケーションを主に行うペアは対象から除外し、手話によるコミュニケーションを主として行うペアのみを被験者として選定した。また、プロトタイプを使用することにより 2.1 節で述べた制約①を取り除くことができるかどうかを検証するため、被験者らを互いの手話が見えにくい座り方で座らせた (図 6)。

そして、プロトタイプを使用してテレビ番組を 40 分視聴するタスク (以下では、プロトありとする) とプロトタイプを使用せず通常のテレビ視聴と同じ状態でテレビ番組を 40 分視聴するタスク (以下では、プロトなしとする) を課した。両タスクを実施する順序は組ごとに入れ替え、それぞれのタスク実施中 20 分が経過したところで、座る位置を交替させた。映像視聴中は普段のとおり自由に会話させ、プロトありで視聴を開始してもらう前には、1~2 分ほど互いのユーザ映像を通して手話をする練習のための時間をとった。PC の映像の出力には、予備検討時と同じ 37 型

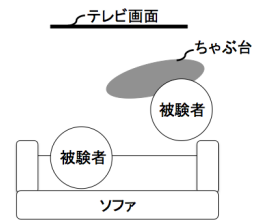


図 6 実験時の座り方を表した模式図
Fig. 6 The diagram shows how the subjects sit.

フル HD 液晶テレビを使用した。視聴してもらう番組映像は、会話が発生しやすいと考えられるバラエティ番組を使用した。プロトありとプロトなしとで同じ番組を使用すると、後で実施したタスク時に番組に飽きた状態で視聴させることになる。番組に飽きることで、プロトタイプの効果を的確にはかることができなくなるのを避けるため、プロトありとプロトなしとは字幕の数が同等である異なる番組を使用した。なお、いずれも被験者に初めて見る映像あることを確認したうえで視聴させた。

またユーザ映像は手話モードを用い、そのサイズは約 17cm × 15cm に固定し、被験者の身長や体格にあわせて微調整を行った。ユーザ映像の表示位置は、4.1 節で述べた予備検討時の要求に基づき、プロトあり実施前に映像下部と映像中央部のどちらかを選択させた。映像視聴中は、各組のコミュニケーションの様子の観察や会話内容の記録を行い、両タスクの終了後には 1 人ずつ、インタビュー調査を行った。

5.3 結果

映像視聴時の記録から会話の合計数と会話の合計の長さを抽出した。

今回の実験結果では、被験者同士の手話での相互のコミュニケーションだけでなく、独り言、笑い声なども会話と見なしている。また、以下の状態が起きた場合、1 回の会話として区切っている。

- 発話をしてから 1 秒以上の間が空いている場合
- 相手が受け答えを行った場合

また、実験中、会話の大部分は手話で行われており、一部確認された口話での会話は感嘆詞を表すものにとどまった。

プロトあり時の会話の合計数からプロトなし時の会話の合計数を差し引いた値を被験者ごとにグラフにして示した (図 7)。同じように、会話の合計の長さについても、プロトあり時の値からプロトなし時の値を差し引いてグラフに示した (図 8)。アルファベットは組を表し、同じアルファベットを持つ 2 名の被験者は 1 つの組をなしている (たとえば、A 組は被験者 A1 と被験者 A2 からなる)。

図 7 および図 8 からは、5 組中 4 組の被験者がちゃぶ台に座ったときとソファに座ったときのどちらかもしくは両方で、プロトあり時の会話数および会話の合計の長さがプ

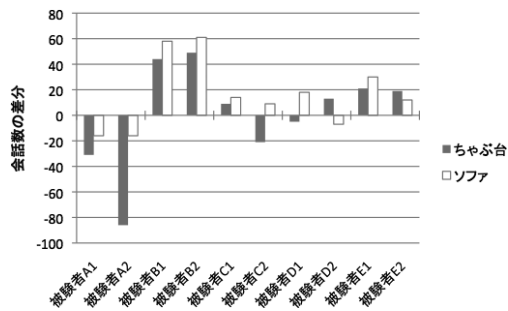


図 7 会話数の差分

Fig. 7 Difference between the conversation frequencies.

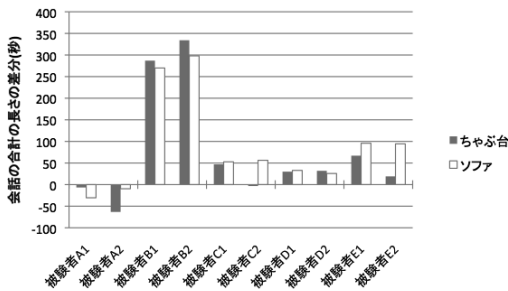


図 8 会話の合計の長さの差分

Fig. 8 Difference between total duration of conversation.

ロトなし時を上回っていることが分かる。一方で A 組のみは会話数・会話の長さともに、プロトなし時がプロトあり時を上回っている。インタビューからはプロトあり時に使用した番組が被験者 A2 の好みに合わなかったことが分かっており、このことが原因で A 組の会話数と会話の長さは、プロトなし時がプロトあり時を上回ったと考えられる。なお、A2 以外の被験者については、視聴した映像が好みに合っていたことをインタビューにより確認している。この結果から、手話モードでプロトタイプを使用したことによって、手話を主に行うコミュニケーションの場合、2.1 節で述べた制約②が軽減されたといえる。

また、この 4 組については、図 6 に示した座り方でも番組映像を見ながら手話モードでのユーザ映像を通して会話を行っていた様子が観察できたことから、プロトタイプを使用することにより制約①が取り除かれたといえる。

6. 考察

以下ではまず、プロトあり時の会話数および会話の合計の長さがプロトなし時を上回った 4 組の被験者が実際に、テレビを視聴しながらも、相手に知ってほしい、共感してほしい、伝えたい、伝えなければならないと思うことを聴覚障がい者特有の制約を受けずに伝えられる状態となっていたかどうか、会話数・会話の長さともにプロトあり時がプロトなし時を最も大きく上回っている B 組を例にとって検証する。まず B 組のインタビューでの回答内容の一部を下記に示す。B 組は聴覚障がい者同士の夫婦で B1 が妻、B2 が夫、Q は質問者を表している。

Q:「お二人の顔を表示して見ていただいたときには、わりと番組の間も会話をたくさんしていただいたと思うんですけども。」

B2:「そうですね。」

Q:「それはかなり意識的に、会話をしようと思ってしていただいた感じですか？」

B2:「そうではなくて、そのときは横じゃなくて正面にいますから話をしたいという気持ちになったんです。」

Q:「やはり、お互いの顔がテレビの画面で確認できることで番組映像も見られるし、お互いも見られるしということで会話をしていただいたということですか？」

B2:「そうですね。はい。ユーザ映像と番組と、一緒ですよ。もしユーザ映像がなければ、やはり顔をここで座って対面で会話をすると見落としが多いので、あまりしないです。」

被験者 B2 の話からは、ユーザ映像が表示されていたことで番組本編中に会話をしても番組を見落とす心配がないという安心感が生まれ、会話することに積極的になったことが分かる。なお B 組の夫婦は普段テレビを視聴する際、番組本編中はあまり会話をせず CM 中や番組終了後に会話をしていることもインタビューから分かった。

さらに、B 組の実験中の会話内容を見てみると、番組内容をきっかけとして会話を始め、会話の最中の番組内容も加味しながら次々に発話を行っていることが分かる。B 組の会話内容の一部を下記に示す。

B1:「(テレビ映像を見ながら) 車が色違いで面白いね」

B2:「面白いね、演出がうまいな」

B1:「そうね」

B2:「1 つに (映像が 1 つになる様子と同期しながら)、最後 1 つになったな」

B1:「うん」

B 組の夫婦は上記の部分以外にも、会話の最中の番組内容を加味した会話を多く行っている。ユーザ映像により番組を視聴しながらの会話が可能になったことで、番組の内容を加味して会話を長く続けるようになったと考えられる。

以上のことから、プロトタイプによりテレビを視聴しながらも、相手に知ってほしい、共感してほしい、伝えたい、伝えなければならないと思うことを聴覚障がい者特有の制約を受けずに伝えられる状態が実現されたといえる。

次に、会話数・会話の長さともにプロトなし時がプロトあり時を大きく上回った A 組について、その要因を検証する。A 組の夫婦はプロトありの映像視聴時においてもほとんどユーザ映像を介さずに、ちゃぶ台前に座っている方の被験者が後ろを向く形で対面によるコミュニケーションを行っている様子が見られた。インタビューではその理由に

ついて回答が得られた。一部を下記に示す。A組は聴覚障がい者同士の夫婦でA1が妻、A2が夫である。

A1:「やっぱり、視線をきちんと合わせて話すというのが習慣づいていますので、ここだと目と目が合わないのちょっと違和感がありました。慣れなかったです。」

A2:「視線が合わなかったのでもずっとそこで会話をする気にはなれなくて。実際に会話したいときには、やはり画面よりも実際に目と目を合わせて話しする方がいいです。」

上記の回答内容からは、被験者A1、A2ともに対面でのコミュニケーションを重視しており、特にユーザ映像を介し視線が合わない状態でコミュニケーションを行うことに強い抵抗感を持っていたことが分かる。インタビューからはさらに、被験者A1、A2ともに互いの手話が見えにくい座り方での映像視聴にも強い抵抗感を持っていたことも分かった。以上の要因が重なり、A組の会話数と会話の合計の長さは、プロトなし時がプロトあり時を大きく上回ったと考えられる。以下では、今後の課題をあげる。

インタビュー調査では複数の被験者から、ユーザ映像が映像視聴時常に表示されているのは、映像視聴の妨げになるため会話を行うときのみ表示されるのが望ましいという指摘があった。今後、一緒に見ている人の顔が見えるテレビインタフェース使用時のユーザエクスペリエンスをより高めていくためには、視点を増やすだけでなく、視点を増やすタイミングを適切に制御することが必要となる。

さらに、被験者Aのようなユーザを含めより多くのユーザに受け入れられるシステムとなるため、相手のユーザ映像を見たときに自分と目が合っている感覚をユーザに持たせる機能を実装していくことも今後の課題である。

また、これら課題は、今回実験を行わなかった口話でのコミュニケーションを対象とした口話モードでのユーザ映像に対してもいえると考えられる。さらに、口話モード実験時の課題となる会話内容把握の正確性担保のために、今後は被験者自身に会話内容を確認させるなど、口話モードにおいても実験を行う。口話と手話をともに使用する場には口話モードと手話モードの併用利用が必要になると考えており、これらについても今後実験を行う。

7. まとめ

本稿では、聴覚に障がいがあるユーザのテレビ視聴時の視点を増やす、一緒に見ている人の顔が見えるテレビインタフェースを提案した。プロトタイプを使用した映像視聴実験からは、プロトタイプを使用しないと比べて、発話数が増える、発話の長さが長くなるなどの結果が得られた。この結果は、インタフェースがテレビ視聴時における

聴覚障がい者特有の制約が軽減されたことを示している。

今後は、ユーザ映像の表示タイミングを制御する機能や、相手のユーザ映像と視線が合う感覚を与える機能を実装することで、本インタフェース使用時のユーザエクスペリエンスをより高めていくことが課題となる。

参考文献

- [1] 本間真一, 小林彰夫, 奥 貴裕, 佐藤庄衛, 今井 亨, 都木徹: ダイレクト方式とリスピーク方式の音声認識を併用したリアルタイム字幕制作システム, 映像情報メディア学会誌, 映像情報メディア, Vol.63, No.3, pp.331-338 (2009).
- [2] 今井 亨: リアルタイム字幕放送のための音声認識, 電子情報通信学会技術研究報告. SP, 音声, Vol.109, No.259, pp.19-24 (2009).
- [3] 総務省: 平成 23 年度の字幕放送等の実績 (2010), 入手先 (http://www.soumu.go.jp/menu_news/s-news/01ryutsu09_02000045.html) (同様のデータを過去 10 年分参照した)
- [4] 平尾美佐, 宮崎泰彦, 東野 豪: 聴覚障がい者と聴者の困らん視聴を支援する映像視聴インタフェースに関する研究, 電子情報通信学会技術研究報告. WIT, 福祉情報工学, Vol.111, No.174, pp.49-53 (2011).
- [5] 志田和也, 長嶋祐二, 本田 学: 手話理解における表情の機能分析方法の基礎的検討, 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理, Vol.108, No.489, pp.7-10 (2009).
- [6] 米原裕貴, 長嶋祐二: 手話の習熟度による注視点の変化に関する実験的検討, ヒューマンインタフェースシンポジウム 2002 論文集 (2002).
- [7] 中園 薫, 米原裕貴, 長嶋祐二, 市川 薫: 手話動画像の評価実験: 画面サイズ, フレームレート, 量子化幅等の影響, 電子情報通信学会技術研究報告. WIT, 福祉情報工学, Vol.105, No.67, pp.19-24 (2005).
- [8] Takeuchi, Y., Sakashita, Y., Wakatsuki, D., Minagawa, H. and Ohnishi, N.: Communication Supporting System in a Classroom Environment for the Hearing Impaired, *ICCHP 2006*, LNCS 4061, pp.627-634 (2006).



平尾 美佐 (正会員)

2008 年慶應義塾大学総合政策学部卒業。2010 年同大学大学院政策・メディア研究科修士課程修了。同年日本電信電話(株)入社。NTT サービスエボリューション研究所にてヒューマンインタラクションに関する研究に従事。



石井 陽子 (正会員)

2004年電気通信大学大学院情報システム学研究科修士課程修了。同年日本電信電話(株)入社。NTTサービスエボリューション研究所にてヒューマンインタラクションに関する研究に従事。2009年電気通信大学大学院情報システム学研究科博士課程修了。平成17年度電子情報通信学会学術奨励賞受賞。Prix Ars Electronica 2008 Interactive Art部門入選。電子情報通信学会会員。



宮崎 泰彦 (正会員)

1989年東京大学理学部数学科卒業。同年NTT入社。1995~1996年まで、米国イリノイ大学大学院計算機学科修士課程修了。主に、設備管理システム、地理情報システム、ITS、Webベースの情報提供システムの研究開発に従事。現在、NTTサービスエボリューション研究所、主幹研究員。



松嶋 信貴 (学生会員)

2012年電気通信大学電気通信学部情報工学科卒業。同年同大学大学院情報理工学研究科博士前期課程進学。現在に至る。ヒューマンコンピュータインタラクション等に興味を持つ。



小林 透 (正会員)

1985年東北大学工学部精密機械工学科卒業。1987年同大学大学院工学研究科修士課程修了。同年NTT入社。以来、ソフトウェア生産技術、ユビキタスコンピューティング、情報セキュリティ等の研究開発に従事。現在、長崎大学工学部工学科教授。電子情報通信学会、IEEE各会員。博士(工学)。