聴覚障害者向けニュースの字幕提示方法に関する主観評価

Subjective Evaluation of Methods for Presenting Closed-Captions on TV News Programs for Hearing-Impaired People

正会員 門馬 隆雄^{*1}, 正会員 沢村 英治^{*2}, 正会員 三橋 哲雄^{*3}
正会員 江原 暉将^{*4}, 白井 克彦^{*1}
Takao Monma^{*1}, Eiji Sawamura^{*2}, Tetsuo Mitsuhashi^{*3},
Terumasa Ehara^{*4} and Katsuhiko Shirai^{*1}

Abstract Several methods for presenting closed-captions (captions) on TV news programs for hearing-impaired people were subjectively evaluated. A total of 180 captioned programs were used for the experiment; produced by combining 36 presentation methods and 5 kinds of TV news materials. The evaluations were had done by 35 hearing-impaired people and 36 people with normal hearing. Three factors were evaluated: overall goodness, caption quality and picture quality. They were evaluated on a scale of 1 to 5. Two-line captions that changed at once were evaluated to be the best. Separating the caption area from the picture area was found to improve the score. Captions were easier to understand when the display time of the captions was lengthened.

キーワード: 主観評価, 聴覚障害者, 字幕提示方法, 字幕評価

1. ま え が き

聴覚障害者向け字幕付きテレビ放送は昭和60年に開始され、すでに14年の歴史がある。しかし、ニュース番組を対象にした同種のサービスは日本においては平成12年に開始されたばかりである。これらの放送の利用者である聴覚障害者はどのような字幕提示方法を好むかを知ること、字幕付きテレビ放送サービスを行う上で重要な要件である。我々は字幕付きテレビ番組の自動制作手法の研究の中で、字幕提示方法の好ましさに対する主観評価実験を行ってきた。聴覚障害者向けの主観評価実験としては初期の予備評価研究¹⁾があるが、観視条件が明らかでなく、この

実験で字幕制作に用いたシステムでは放送品質の字幕番組を制作することができず、また字幕提示方法として文字サイズの影響を検討しておらず、映像の圧縮率や単位時間あたりに表示すべき字幕の量などの評価もしていない。一方、本論文で述べる主観評価実験では放送品質の字幕番組を利用している。主観評価実験の報告²⁾では、字幕番組の制作手法が中心であり、本論文ではそれを発展させ評価実験の分析を詳細に行う。また、実験結果と字幕番組の概要について述べた報告³⁾もあるが、本論文で行っている映像の圧縮率や時間あたりの字幕の読み取り量などについて未評価である。字幕番組制作システムと実験の概要説明と定性的な評価の報告⁴⁾があるが、本論文では定量的な評価に重点をおいた。

文字の識別可能性や画面のコントラストなどの視聴条件、視覚特性などによる画面上の文字表示に関する最適な条件については、文字放送における文字のコントラストなどの評価分析⁵⁾などがあるが、これらの研究では実験対象者が映像のプロである放送技術者であるものが多数を占めてきた。さらに、評価対象もニュース番組に聴覚障害者向け字幕を付与したものではない。筆者らの目的は、ニュース字幕という新しい番組形態に対してエンドユーザである聴覚障害者の好みを知り、それを字幕付きテレビシステム的设计に生かすことにある。この方向の先行研究としては、聴覚障害者を対象とした字幕提示方法の主観評価に関する内容理解度の評価⁶⁾がある。しかし、この実験では、

1999年10月21日、通信学会HCS研究会で発表

2000年2月29日受付、2000年5月12日再受付、2000年7月31日採録

*1 早稲田大学/(TAO)

(〒157-8510 東京都新宿区戸塚町1-104, TEL 03-3203-8714)

*2 通信・放送機構 (TAO)

(〒151-0064 東京都渋谷区上原3-23-5, TEL 03-5790-7065)

*3 尚美学園大学/(TAO)

(〒350-1153 川越市下松原655, TEL 0492-46-5251)

*4 NHK放送技術研究所/(通信・放送機構)

(〒157-8510 東京都世田谷区砧1-10-11, TEL 03-5494-2308)

*1 Waseda University (1-104, Totsuka-cho, Shinzuyku-ku, Tokyo 157-8510, Japan)/TAO

*2 Telecommunications Advancement Organization of Japan (TAO) (3-23-5, Uehara, Shibuya-ku, Tokyo 151-0064, Japan)

*3 Syobi Gakuen College

(655, Shimomatsubama, Kawagoe-shi, Saitama 350-1153, Japan)/TAO

*4 NHK Science and Technical Research Laboratories

(1-10-11, Kinuta, Setagaya-ku, Tokyo, 157-8510, Japan)/TAO

字幕提示方法として、現行の字幕放送に準拠した2行字幕と、字幕を表示するための専用領域を設けた8行字幕の2種類を検討したに過ぎない。したがって、これらの限られた研究から、どのような字幕が好まれるかという字幕放送用システムの具体的な設計指針を示すのは困難である。一方、今後実現されるデジタル放送では、デジタル映像技術の発達に伴い、ユーザの好みに合わせ多様な字幕提示方法を選択可能な字幕受信機や字幕放送方式が求められることが予想される。このような背景から、字幕放送のユーザである聴覚障害者の意見を反映し、聴覚障害者向け字幕の需要の大きな番組であるニュース番組に対する字幕番組制作システムにおいて利用可能な知見を得るため、ニュースの字幕提示方法の主観評価実験を実施した。実験では聴覚障害者団体に御協力をお願いした。以下、2章では実験の内容について、3章では実験結果の分析について、4章で結論を述べる。

2. 実験の内容

2.1 字幕提示パラメータ

実験の目的から、様々な字幕提示方法を用意する必要があるが、主観評価の対象とした字幕提示パラメータとその内容を表1に示す。これらの4パラメータについて、それぞれの組合せを評価した。詳細な説明を(1)～(4)項で行う。

(1) 字幕行数
表示字幕はすべて横書きとし、標準的な2行字幕を中心に、1行および3行を加えた3種類を対象とした。

(2) 文字サイズ
モニタ上で現行の字幕放送の漢字フォントとほぼ同一サイズとなるCRT上で縦横64 dotの正方形フォントを「標準」と呼び、縦を横の2/3倍の長さに縮小した横長のフォントを「縦2/3」、横を縦の2/3倍の長さに縮小した縦長のフォントを「横2/3」と呼ぶことにする。字間・行間の余白を含めると、モニタの1画面に、標準サイズの文字は横16文字、縦2/3の文字は横22文字入る。

(3) 字幕領域
これは字幕を表示するときに字幕表示専用の領域をモニタ画面上に用意するかどうかを制御するパラメータである。この領域を「字幕領域」と呼ぶ。字幕領域がある場合、映像を縦横に縮小することによって生じた背景が黒の

領域に字幕を表示する。字幕領域がなしの場合、映像は非縮小で映像と字幕とを重畳させる。字幕領域がある場合、映像の縮小の割合を映像の圧縮率と呼ぶ。

(4) 字幕転換方法
字幕転換方法は、行数が1行のとき総入替および横スクロールとし、2行および3行のとき総入替および縦スクロールという、3種類になる。

横スクロールの場合の字幕の行数は、以下で理由を説明するが1行のみとした。電光掲示板のように字幕の最初の1行を静止状態で提示し、最初に表示した字幕文を読み終わった頃合いから、番組の字幕の文字数に応じた一定速度で右から左へ文字が移動する。番組終了より若干前のタイミングでスクロールが停止し、字幕文章は文末を表示したまま停止する。このためニュース映像と字幕の内容は必ずしも同期していない。

一方、縦スクロールは、文字が下から上にスムーズなせり上がりとし、静止を断続的に繰り返しつつ表示する。これはアメリカで実施されている字幕放送に似せた字幕提示方法であるが、実際のアメリカの字幕放送では単語単位に字幕文が表示されるのに対し、今回の実験では文単位で表示される点が異なる。せり上がり時間は200 msで、各字幕が静止している時間は各字幕の表示タイミングに応じて異なる。現在、アメリカの字幕放送では、字幕はリアルタイムで手動で入力されるのが一般的であり、自動化されてはならず、発話が早過ぎる場合には音声に対して字幕が追いつかない場合も多い。一方、この実験では常に字幕と音声と同時に終了するようにした。

(5) パラメータの組合せ数と例
パラメータの種類は、行数が3種類、文字サイズが3種類、字幕領域が2種類、字幕転換方法が3種類となり、これらの組合せは54種類となる。しかし、この実験では以下の想定から行数と字幕転換方法を合わせて6種類に制限した。このためパラメータの組合せは36種類である。

- ・横スクロールで行数が1行以外の場合、例えば2行横スクロールなどは実際の放送でも存在しているが、ドラマ番組の出演者名の表示などであり、ニュース番組では例がなかったため、この実験では検討から外した。
- ・縦スクロールで字幕行数が1行の場合、総入替形で字幕行数が1行のものと区別がつかないので、これも検討から外した。

図1～図4に作成した字幕番組の例を示す。

2.2 字幕表示方法の詳細

字幕はパソコンに内蔵されているフォント(MSゴシック)を業務用ダウンキャンコンバータでNTSC信号に変換し親映像と合成した。映像の圧縮率は画面下部の字幕領域と親映像が重畳しないように決定した。文字サイズが縦2/3で行数が3行のときの映像の圧縮率は、文字サイズが標準で行数が2行のときと同等である。文字サイズが横

表 1 字幕提示パラメータの一覧
List of presenting methods parameters.

	方法 1	方法 2	方法 3
字幕行数	1 行	2 行	3 行
文字サイズ	標準	縦 2/3	横 2/3
字幕領域	あり	なし	
字幕転換方法	総入替	横スクロール	縦スクロール

図 1 2行・標準・字幕領域ありの字幕の例
The example of the caption of 2 lines, standard font, with caption area.

図 3 1行・横2/3・字幕領域ありの字幕の例
The example of the caption of 1 lines, side 2/3 font, with caption area.

図 2 3行・縦2/3・字幕領域ありの字幕の例
The example of the caption of 3 lines, length 2/3 font, with caption area.

図 4 2行・標準・字幕領域なしの字幕の例(現行字幕放送)
The example (in the same way as the present closed-caption broadcasting) of the caption of 2 lines, standard font, without caption area.

2/3のときの映像の圧縮率は、文字サイズが標準のときと同等である。実際にビデオを収録するときに、以上の映像の圧縮率に加え、テレビの有効画素範囲であるセーフティゾーンを考慮した85%の画像圧縮を行っている。文字サイズが標準で行数が1~3行のときの映像の圧縮率は、それぞれ70%, 64%, 57%となり、文字サイズが縦2/3ではそれぞれ74%, 68%, 64%となる。

2.3 評価項目

この実験では字幕番組の「全体の良さ」に加え、字幕番組にとって最も重要である「映像(親映像)の見やすさ」と「字幕の見やすさ」を評価項目とした。予備評価実験¹⁾ではそれに加え「字幕のわかりやすさ」を評価項目に加えたが、字幕の見やすさとほとんど同一の結果しか得られず、今回の評価実験では、映像と字幕の評価、さらに字幕番組全体の良さの評価への影響にポイントを絞ることにした。評価項目を限定することにより、評価に要する時間を短くすることができ、その結果、字幕提示方法を多様にすることができた。また、評定者が映像技術者でなく映像評価と

しては素人であることなどから、評価項目は簡素であることが望ましいと考えた。以上から、字幕提示方法の評価項目としてはこれらの3項目で充分であると考えた。

本実験では、オープンキャプションと呼ばれるテレビ放送の映像内に元々入っている字幕の見やすさは、映像の見やすさとして評価するようにした。一方、クローズドキャプションと呼ばれる、字幕放送デコーダによって表示される字幕の見やすさを、字幕の見やすさとして評価するようにした。実験の開始前には、デモ映像とOHPに加え、聴覚障害者には手話を併用して誤解がないように説明した。

実験は、評価項目として、全体の良さ、映像の見やすさ、字幕の見やすさの3項目に、評点5水準(5 良い, 4 やや良い, 3 普通, 2 やや悪い, 1 悪い)で回答することとし、また各字幕番組に対して自由記述で意見を述べるようにした。

2.4 観視条件

この実験では、すべての観視者はニュース番組の音声か

すべて消去された状態で観視した。実験における観視条件はITU-R BT. 500-7に準ずるように設定した。1台のモニタを観視する評定者の人数は3人とし、評定者はモニタから6Hの距離に30°以内の角度で配置した。実験の様子を図5に示す。

2.5 ニュース番組素材と字幕素材

本実験では、日本テレビ放送網(株)が放映したニュース番組(番組名:ニュースダッシュ)から、政治・経済・文化・事件・スポーツの5分野のニュース番組を評価素材として選択し、各素材の放送時間が1分未満となるように映像編集を行った。表2に使用した素材の内容を示す。各素材の時間を秒で示した。字幕文章は書き起こしたニュース原稿に対して手作業の要約で作成した。要約作業では、音声から書き起こした原稿の文字数に対して72~74%の範囲になるように文字数を削減したが、1文全体の削除はし

図5 実験の様子
The evaluation experiments.

表2 評価番組用ニュース素材
News materials for the evaluation program.

種目	放映日	時間 (s)	要約率	字幕字数	字幕行数 (A)	(B)
政治	1997/11/19	53.30	0.733	215	16	14
経済	1997/11/13	56.07	0.734	267	20	16
文化	1997/11/18	48.70	0.730	216	16	12
事件	1997/12/9	47.53	0.726	210	15	12
スポーツ	1997/12/9	51.37	0.727	259	20	16

ないようにして、この要約によって音声と字幕のタイミングが大幅にずれることがないようにした。以下、字幕素材の作成方針の詳細について説明する。

(1) 字幕の表示消去タイミング

各字幕の表示タイミングは、字幕と対応する音声の開始と終了が、字幕の表示開始と表示終了に同期するように1/30秒単位で設定した。総入替形字幕のタイミングについては0.5秒以上の無音声区間の前後で字幕の入替があるとき、無音声区間では厳密に字幕を消去することとした。このことにより字幕の行数が1行のときは無音声区間の判定による小刻みな字幕の消去が行われるが、2行や3行表示では字幕文中に無音声区間が含まれる場合が1行に比べ多くなる。

(2) 字幕の「1秒あたりの読み数」

1秒間にどの程度の字幕を表示するか決定することは、自動字幕制作システムでは重要な課題である。本論文では、字幕を仮名に直したときの仮名数を「読み数」と呼ぶ。字幕が表示されてから消去されるまでに経過する秒数で読み数を除算して「1秒あたりの読み数」が求まるが、これは字幕番組中の字幕文、字幕行数や文字サイズなどの字幕提示方法に依存して変化する。(1)に示した字幕の表示消去タイミングの厳密性により、1秒あたりの読み数は1行のときが最も多く、2行や3行のときはそれよりも少なくなる。

(3) 字幕番組の行数と1行あたりの最大表示文字数

標準サイズ、横2/3サイズの文字を用いたときの字幕の行数は、表2でそれぞれ行数(A)、(B)で示した。縦2/3サイズでは字幕の行数は標準サイズと同等である。標準サイズでは1行あたりの文字数は通常16文字以内である。ただし、文末文字が句点および読点の場合、1行に17文字表示する。横2/3サイズでは1行あたりの最大文字数は22文字である。

(4) 制作した全字幕番組数

5種類の映像素材に対し36種類の字幕提示方法があるので、字幕つきニュース番組は総計180本である。

2.6 評定者

以下の3評定者グループによって観視と評価が行われた。各評定者の矯正視力は1.0以上である。また、級とは聴覚障害者の障害者手帳にあり、級が少ないほど重い障害であることを示している。

- (1) 聴覚障害者 中途失聴者18名(2級14人、3級2人、4級2人)、平均年齢49、最年長69、最年少36(全日本難聴者・中途失聴者団体連合会、東京都中途失聴・難聴者協会)
- (2) 聴覚障害者 聾者17名(1級5人、2級12人)、平均年齢39、最年長63、最年少28(東京都聴覚障害者連盟)
- (3) 健聴者36名、平均年齢23、最年長45、最年少19(学生31名、主婦5名)

2.7 実験の実施

1番組について、評定者はそれぞれ約1分間、素材の観視と評価用紙の記述を交互に行った。1回の実験で1人の評定者は60番組を観視し、15番組毎に40または45分の休憩を行った。また、各字幕提示方法の提示順序については乱数を用いて学習効果がないように決定された。

3. 実験の結果

本章では実験の結果について述べる。まず評価項目間の相関関係を分析した。映像の見やすさと字幕の見やすさの相関係数は0.45752となり中程度の相関があった。全体の良さと字幕の見やすさの相関係数は0.73620、全体の良さと映像の見やすさの相関係数は0.74844となり、それぞれ強い相関があった。

回帰係数を用いて各評価項目間の関係を示すと以下のようになる。全体の良さは字幕の見やすさと映像の見やすさをバランス良く同程度に反映することが示された。

$$\begin{aligned} (\text{全体の良さ}) = & 0.157 + 0.469 \times (\text{字幕の見やすさ}) \\ & + 0.476 \times (\text{映像の見やすさ}) \end{aligned}$$

上記の結果を踏まえ、3.1節から3.5節まで全体の良さのみに注目して詳細に分析を行った。全体の良さと映像の見やすさ、字幕の見やすさの分析を3.6節で行った。実験結果をより詳細に分析するため、各評価項目と、映像の圧

縮率、1秒あたりの読み数の関係の分析を、それぞれ3.7節と3.8節で実施した。

3.1 字幕提示方法と全体の良さの評価値の分布

全体の良さの評価値について、図6で聴覚障害者と健聴者を合わせた評定者全体の評価値順に、字幕提示方法と評定者全体の評価値分布、聴覚障害者と健聴者における順位、聴覚障害者と健聴者の評価値に有意差があったかを示す。平均評価値は、最上位の3.97から最下位の2.68の間に分布している。現行字幕放送での字幕提示方法は字幕領域なしの中では最上位で17位となった。一方、上位16位までを字幕領域ありのものが占めた。最上位の評価の字幕提示方法は現行の字幕放送の字幕提示方法で字幕領域をありとしたものとなった。

聴覚障害者と健聴者の評価値の差異について分析すると、聴覚障害者で1位の字幕提示方法は評定者全体の1位に一致した。しかし、健聴者の1位は聴覚障害者による評価では10位の字幕提示方法となった。次に、聴覚障害者と健聴者の各グループ間で評価値に危険率5%で有意差があったものを調べたが、4種類の字幕提示方法では聴覚障害者の評価は健聴者の評価に比べ、有意に高かった。表3に有意差があり、特に聴覚障害者に好まれた字幕提示方法を示す。一方、残り32種類の字幕提示方法には有意差がなく、したがって、本実験からは聴覚障害者と健聴者で字

図6 全体の良さについての字幕提示方法の順位

Evaluated score for the caption presentation methods measured by "the overall goodness".

幕提示方法の好ましさの順位の傾向に違いは少ないことがわかった。ここで有意差があった字幕提示方法を調べると、4種類中3種類が3行字幕で、同じく3種類が総入替であった。

一方、個々のパラメータについて調査すると、字幕行数では3行、文字サイズでは標準、字幕転換方法では総入替、字幕領域の有無については、ありのとき、聴覚障害者の評価値は健聴者に比べ高く、危険率5%で有意差があった。

全種類の字幕提示方法についての平均評価値を聴覚障害者と健聴者で比較したところ、聴覚障害者の評価値が健聴者の評価値に比べて高く、危険率5%で有意差があった。好ましさの順位の傾向については聴覚障害者と健聴者で差異はないが、全体としては聴覚障害者の評価は健聴者に比べて高く、字幕放送の拡充に対しての期待度の違いが評価値の差の原因のひとつとなったのではないかと考えられる。

3.2 全体の良さの上位の提示形式についての考察

図6で、順位1〜7位のグループにおける全体の良さに対する評価値は、10位以下のグループの評価値との間に危険率5%で有意差があった。上位1〜7位の字幕の主な特徴と考察は以下の通りである。

- すべて字幕領域ありの字幕形式であり、字幕領域ありの有効性が示された。

表 3 特に聴覚障害者に好まれる字幕提示方法
Presenting methods liked specially by the hearing impaired person.

行数	文字サイズ	字幕転換方法	字幕領域	全体順位	聴覚障害者順位	健聴者順位
3	縦 2/3	総入替	あり	2	2	7
1	縦 2/3	総入替	あり	10	7	18
3	横 2/3	総入替	あり	13	11	21
3	標準	総スクロール	あり	21	13	32

- 字幕転換方法では、上位7種類中6種類が総入替でありスクロールより総入替が好まれることがわかった。
- 字幕行数と文字サイズに関しては、1行では画面内の同時表示文字数が多くなる横2/3、3行では映像が大きくなる縦2/3が好まれることがわかった。

3.3 全体の良さの下位の提示形式についての考察

図6で下位の評価が与えられた字幕提示方法について、主な特徴と考察は以下の通りである。

- 全体順位の23位以下のグループは、すべて字幕専用領域なしの字幕形式である。親映像との重畳の結果評価値が低下したと考えられる。
- 28位以下のグループは、スクロールか、1または3行の字幕となった。
- スクロールは評価値が低かった。

3.4 全体の良さと字幕パラメータの分析

本章では3.2節、3.3節で行った考察を裏付けるために統計的な分析を行う。まず、字幕パラメータがどのように字幕番組評価値と関わるかを分析するために、各パラメータの多重比較分析を行った。特に、全データと字幕領域ありの場合に関して、全体の良さと字幕パラメータの関係を表4に示した。

文献7)などによれば、一般にt検定は2つのグループの平均の違いを示すのに有効な手法であるが、グループの個数がNになったときに $N \times (N - 1) / 2$ (個)のペアに対してt検定を行うことは繁雑であり、結果の解釈も困難がある。そこで、グループが3つ以上の場合に各グループの母平均に違いがあるかどうかを分散の大きさの違いで検定するときに有効な手法が分散分析である。このとき、全体の中でどこどこに差があるかを示すときに行うのが多重比較分析という手法である。

有意差のあった*で示されるパラメータについて説明する。字幕領域は他のパラメータとはかなり異なった性質のものであると考えられるが、字幕領域ありと全データとの

表 4 全体の良さに対する字幕パラメータの多重比較分析
The multiplex comparative analysis of the caption parameters to the score of “the overall goodness”.

	df	全データ				字幕領域あり			
		MS				MS			
	効果	効果	F	p-値		効果	F	p-値	
行数	2	11.671	12.825	3 E-06	*	12.085	13.209	2 E-06	*
文字サイズ	2	2.2017	2.4195	0.0891		1.212	1.3248	0.2661	
字幕転換	1	54.996	60.437	1 E-14	*	55.026	60.145	1 E-14	*
字幕領域	1	350.45	385.12	0	*				
行数・サイズ	4	14.797	16.261	3 E-13	*	9.4021	10.277	3 E-08	*
行数・転換	2	14.621	16.067	1 E-07	*	9.1159	9.9639	5 E-05	*
サイズ・転換	2	4.0262	4.4245	0.012	*	1.1559	1.2634	0.2829	
行数・領域	2	2.4796	2.7249	0.0657					
サイズ・領域	2	0.1494	0.1642	0.8486					
転換・領域	1	10.063	11.058	0.0009	*				

相違点としては、評価値が全データに対しに 0.3 程度高いこと、文字サイズ・字幕転換方法に有意差がないことがある。字幕領域ありにより評価が良くなったことで他のパラメータの及ぼす影響が小さくなったものと考えられる。

表 4 で有意差のあることがわかった字幕パラメータについて、全体の評価値と個々の字幕パラメータに対する関係を説明する。図 7 に示す通り、字幕行数では、3 行が 1 行

図 7 字幕行数と全体の良さ
Evaluated “the overall goodness” score with the number of caption lines.

図 8 字幕転換方法と全体の良さ
Evaluated “the overall goodness” score with caption-changing methods.

図 9 字幕領域と全体の良さ
Evaluated “the overall goodness” score with the caption area.

と 2 行に対し 0.15 程度低く、有意差があった。図 8 に示す通り、字幕転換方法では、総入替に比べスクロールが 0.22 程度低く、有意差があった。図 9 に示す通り、字幕領域では、字幕領域ありは字幕領域なしに比べ 0.6 程度と大幅に良くなり、有意差があった。

3.5 複合パラメータの分析

本章では字幕パラメータのいくつかの組合せについて分析した。

(1) 字幕領域と字幕転換方法

字幕領域と字幕転換方法についての関係を調査し、図 10 に示す。字幕領域に関わりなく総入替がスクロールに対し評価値が高く有意差があり、字幕転換方法に関わりなく字幕領域ありが字幕領域なしに対し評価値が高く有意差があった。

(2) 字幕行数と文字サイズ

字幕行数と文字サイズの組合せを図 11 に示した。縦 2/3 では行数が増加すると評価が良く、標準および横 2/3 では逆である。1 行では、標準および横 2/3 がともに同程度に良く、縦 2/3 が悪くなり、縦 2/3 のみ有意差があった。一方、3 行では縦 2/3 が良い評価であり、標準、横 2/3 の順となっていてそれぞれ有意差が認められた。2 行の場合は、標準、横 2/3、縦 2/3 の順であるが、標準と縦 2/3 の間にのみ有意差が認められる。文字サイズについては単独

図 10 字幕領域と字幕転換方法
The caption area and caption-changing methods.

図 11 字幕行数と文字サイズ
The number of lines and character size.

図 12 字幕転換方法と字幕行数
Caption-changing methods and the number of lines.

では効果が小さかったが、文字サイズと他のパラメータとの組合せによる効果は大きかったことがわかった。

(3) 字幕転換方法と行数

字幕転換方法と字幕行数の組合せを図 12 に示した。行数から見た総入替とスクロールの比較では、1 行ではほぼ同じ評価値であり、2 行と 3 行では総入替が高い値となり有意差があった。字幕転換方法から見た行数との比較では、総入替では 1 行と 3 行で評価値がほぼ等しいのに対し 2 行では高く有意差があったが、スクロール形字幕転換方法では、縦スクロールである 2 行と 3 行に比較して横スクロールである 1 行が高く有意差があった。

3.6 評価項目の評価値と字幕パラメータの分析

主観評価実験結果をもとに字幕の各パラメータと各評価項目ごとの関係について定量的な分析を行う。ここで目的変数である評価値は 5 水準であり量的変数であるが、説明変数である字幕パラメータは質的変数である。以下では数量化理論 I 類を適用することとし、説明変数である各字幕パラメータには現行の字幕放送に近い値を基準としてダミー変数を割り当てることにした。現行字幕表示形式は横 2 行の字幕を総入替え形の字幕転換方法で映像に重ねて表示しているが、この表示形式を比較対象として、各評価項目を従属変数、各字幕パラメータを独立変数として重回帰分析を行い、表 5 に示した。以下で、ダミー変数 LINE 1 (1 行)、LINE 3 (3 行)、CHAR 1 (文字サイズ縦 2/3)、CHAR 2 (文字サイズ横 2/3)、SCROLL (スクロール)、AREA (映像圧縮あり) に対しての数値の割り当ては、括弧内に示す条件が満たされた場合に 1 とし、そうでなければ 0 とした。このようにしてそれぞれの変数と評価項目ごとの相関を調べ、評価項目を字幕パラメータの関数として定量的に示した。各数式で、下線付きの字幕パラメータは有意差がある。

$$(\text{全体の良さ}) = 3.151 + 0.038 \text{ LINE } 1 - 0.135 \text{ LINE } 3 - 0.082 \text{ CHAR } 1 - 0.049 \text{ CHAR } 2 - 0.226 \text{ SCROLL} + 0.594 \text{ AREA}$$

全体の良さの評価値に対し、字幕領域ありは最も良い影響を与えるが、スクロールは最も悪い影響を与える。ま

表 5 各評価と字幕パラメータの重回帰分布
The evaluated scores and the caption parameters.

	全体の良さ		字幕の見やすさ		映像の見やすさ	
	B	p-値	B	p-値	B	p-値
Y 切片	3.1508	0	3.2470	0	3.1900	0
LINE 1	0.0377	0.312	-0.0270	0.507	0.1930	5 E-07
LINE 3	-0.1350	3 E-04	0.1065	0.009	-0.330	1 E-17
CHAR 1	-0.0820	0.028	-0.3090	3 E-14	0.1060	0.005
CHAR 2	-0.0490	0.191	-0.0540	0.189	-0.0800	0.032
SCROLL	-0.2260	2 E-13	-0.3790	8 E-30	-0.1000	8 E-04
AREA	0.5940	0	0.5510	0	0.5980	0

た、3 行字幕と縦 2/3 の文字サイズは悪影響である。

$$(\text{字幕の見やすさ}) = 3.245 - 0.027 \text{ LINE } 1 + 0.106 \text{ LINE } 3 - 0.309 \text{ CHAR } 1 - 0.054 \text{ CHAR } 2 - 0.379 \text{ SCROLL} + 0.551 \text{ AREA}$$

字幕の見やすさの評価値に対し、字幕領域ありと 3 行字幕は良い影響を与えるが、スクロールと縦 2/3 の文字サイズが悪い影響を与える。

$$(\text{映像の見やすさ}) = 3.190 + 0.193 \text{ LINE } 1 - 0.331 \text{ LINE } 3 + 0.106 \text{ CHAR } 1 - 0.083 \text{ CHAR } 2 - 0.105 \text{ SCROLL} + 0.598 \text{ AREA}$$

映像の見やすさの評価値に対し、字幕領域ありと 1 行字幕、縦 2/3 の文字サイズは良い影響を与え、3 行字幕、横 2/3 の文字サイズ、スクロールが悪い影響を与える。

映像と字幕の評価について要因を分析すると、3 行字幕と縦 2/3 の文字サイズの影響が正反対であることがわかった。行数と文字サイズに依存するパラメータで、字幕制作システムで特に検討すべき事項としては、映像の圧縮率と、字幕の 1 秒あたりの読み数があるが、これらのパラメータが、結果をわかりやすく示すために役立つと考えた。このため、3.7 節では映像の圧縮率について、3.8 節では 1 秒あたりの読み数について、それぞれ各評価項目と分析を行った。

3.7 映像の圧縮率と評価項目との分析

字幕領域ありの字幕提示方法が字幕領域なしの字幕提示方法に対して良い評価を受けたが、字幕領域を設けるためには映像の圧縮が必要である。字幕領域ありの場合の、映像の圧縮率の影響について検討を行った。映像の圧縮率は文字サイズと行数によって決定されるが、映像を縮小することによる映像の見やすさへの影響は大きい。将来の映像圧縮を用いた字幕受信装置やインターネットの動画による字幕番組の放映のためにも、映像の圧縮率に着目した評価は必要であり、また、文字サイズと行数を映像の圧縮率に集約することにより分析がしやすくなると考えた。

表 6 に映像の圧縮率と各評点の平均値を示す。評点 3.5 を越えた場合を好ましいと考えた場合、字幕の見やすさについては映像よりも文字が同時に多く、表示される圧縮率

表 6 映像の圧縮率と評点平均値
The image compression rates and evaluated score.

圧縮率	57%	64%	68%	70%	72%	74%
字幕	3.6651	3.5816	3.2576	3.5938	3.7280	3.0625
映像	3.0860	3.6485	3.8690	4.1150	4.1200	4.0535
全体	3.2990	3.4680	3.5589	3.7718	3.8480	3.4010

表 7 総入替形字幕の映像の圧縮率と評点平均値
The image compression rates of caption changing at once and evaluated score.

圧縮率	57%	64%	68%	70%	74%
字幕	3.9720	3.9755	3.6260	3.6601	3.1120
映像	3.1355	3.7951	3.9512	4.1330	4.0720
全体	3.4486	3.8991	3.7886	3.8276	3.4400

が72%以下の場合好ましいという結果が得られ、映像の見やすさについては、親映像の領域が大きい映像の圧縮率が64%以上の場合に好ましいという結果が得られた。また、全体の良さについては、圧縮率が68%以上72%以下の場合に好ましいという結果が得られた。

さらに、字幕提示方法の影響を除去するため、総入替形の字幕992番組について、同様に評点3.5を好ましいとした場合の検討を行った。表7に映像の圧縮率と総入替形字幕の各評点の平均値を示す。映像の圧縮率と字幕の見やすさの相関係数は-0.226で、弱い負の相関があった。つまり、字幕の表示領域が増加するほど字幕の評価が向上している。圧縮率と映像の見やすさの相関係数は0.339で、弱い正の相関があった。これは映像の縮小はできるだけしない方が、映像の評価が高くなるということを示す。全体の良さが評点3.5を上回る字幕の映像の圧縮率は64~70%であった。この実験の範囲では、映像の圧縮率が64~70%の字幕と、57%以下または74%以上の字幕間で比較すると、両グループに有意差があった。総入替形の字幕では、映像の圧縮率が64~70%の範囲になるよう行数や文字サイズの組合せを選択することが好ましいと結論できる。

3.8 1秒あたりの字幕の読み数と評価項目の分析

字幕の読み取りにかかる時間を分析するため、漢字の影響を考慮し、2.5節(2)で定義した読み数を調査し、評価項目に対しての分析を行った。1秒あたりの読み数は行数などのパラメータによって変わるが、文字サイズが標準、総入替、映像領域ありの字幕に限定すると、最小6.1~最大7.2の範囲である。字幕領域がありの場合に限定し、字幕転換方法を総入替、文字サイズを標準に限定した字幕318種について、1秒あたりの読み数と字幕・映像・全体の評点の平均値を表8に示した。

字幕の見やすさと1秒あたりの読み数の相関係数は

表 8 総入替形字幕の1秒あたりの読み数と評点平均値
The number of the readings per one second of caption changing at once and the evaluated score.

読み数/s	6.1	6.3	6.4	6.5	6.6	6.7	7	7.2
字幕	3.954	4.361	4.311	4.500	3.961	3.931	3.121	3.115
映像	3.551	3.583	3.466	3.277	3.654	3.277	3.951	3.782
全体	3.609	3.802	3.822	3.777	3.807	3.653	3.414	3.396

-0.240で負の相関があり、読み取るべき情報に相当する字幕の読み数が減るほど字幕の評価は向上する。映像の見やすさと1秒あたりの読み数の相関係数は0.150であったが、自由記述意見によると、字幕を無視してしまうことにより映像の情報を受け取りやすくなる現象があるという意見が見られた。全体の良さの評点が3.5以上の字幕の1秒あたりの読み数は、6.1~6.7の範囲であった。1秒あたりの読み数が7.0以上の字幕と6.1~6.7の字幕を比較すると、有意差があった。7.0以上になったのは行数が1行の字幕であり、厳密に字幕表示タイミングを音声区間に同期させている。この実験では、音声区間の前後の無音声区間まで含んで字幕を長く表示する1秒あたりの読み数が少ない字幕提示方法が、厳密に無音声区間を判定して字幕の表示タイミングを決定する字幕提示方法に比べ好ましいことがわかった。

4. 結 論

主評価実験の結果から、聴覚障害者向け字幕付きニュース番組の字幕提示方法について以下のような結論を得た。

(1) 主観評価実験により、好ましい字幕提示方法と、その要因を明らかにした。上位の7種類の字幕提示方法とそれ以外には有意差があった。

本実験では、現行の字幕放送と互換性のある行数2行、文字サイズ標準、総入替の字幕提示方法に対して、字幕領域がありの映像圧縮を行ったものが最も好まれており、推奨されることが示された。

(2) 聴覚障害者と健聴者のグループ別の評価では、字幕提示方法の順位に対しての好ましさの差があったのは36種類中4種類だけであり、両グループで好ましさの順位の差は少なかった。一方、字幕提示方法を考慮せずに全体を比較すると、聴覚障害者の評価が健聴者の評価より高く有意差があった。これは自動字幕番組制作システムに対する期待感が、聴覚障害者は健聴者に比べて高いという違いによる影響が大きかったものと推測される。また、聴覚障害者にとって見慣れた現行字幕放送と同じ字幕提示方法が、聴覚障害者に最も高い評価を得ている一方で、聴覚障害者では健聴者よりも3行表示を好ましく感じており、有意差があることがわかった。

(3) 各字幕パラメータについて、単独あるいは組合せによる評価への影響を分析した。総入替はスクロールより

も有意に好まれる。一方、スクロールの種別では、列車内の電光掲示板などで一般的な1行横スクロールが上位に入り、好まれることがわかった。

親映像の圧縮による字幕領域をありとする字幕提示方法が有意に好まれており、特に聴覚障害者では健聴者に比べ有意に好まれ、極めて効果的な字幕提示方法であるといえる。今後の字幕放送において、放送局側か字幕デコーダ側かのいずれかで映像圧縮機能を実装することの有効性が示された。映像圧縮の技術的な実現方法については、デジタル映像処理技術の発達、大容量メモリの低価格化により、現行の字幕受信機の機能として検討すべきであろう。字幕領域ありの字幕提示方法について、早期の実用化が望まれる。

(4) 字幕領域ありの場合について、映像の圧縮率の評価への影響を調べた。特に総入替形の字幕については、映像の圧縮率により生ずる字幕領域の割合に応じて、字幕領域が大きければ字幕の見やすさの評価が向上し、映像領域が大きければ映像の見やすさの評価が向上することが示された。特に総入替では映像の圧縮率と評価項目に対して映像の圧縮率が64~72%の範囲では、全体の良さの評点が3.5を上回り、好まれることがわかった。

放送の内容に応じて受信者の判断で映像と字幕のどちらか一方を優先したいという要求は自然なものである。そのためには、字幕提示方法を受信者側で選択できる字幕受信機が有効であると考えられる。

(5) 総入替形字幕の制作には、字幕をどのタイミングでどの量だけ表示するかに対する基準が必要であるが、字幕中の読み仮名に対応する読み数が全体の良さの評価と関連があるかを調べた。特に文字サイズが標準のとき、1秒あたりの読み数が6.1~6.7の範囲の、字幕の表示タイミングに音声区間の前後の無音声区間を含むような字幕提示方法が好ましいことがわかった。字幕の1秒あたりの読み数を、音声区間の前後の無音声区間を利用して最適化する字幕生成手法は、好ましい字幕番組の制作に有効であろう。字幕文章に対し適切な箇所で行や改ページを挿入し、視聴者の字幕の読み取り時間を考慮して表示タイミングを計算する、自動的な最適な字幕画面制作要素技術の研究において、本実験によって得られた知見は有効に活用できる。

5. む す び

聴覚障害者を含む評定者を対象とした字幕付きニュース番組の字幕提示方法に関する主観評価実験について報告し

た。この主観評価実験で得られた成果を、今後のデジタル放送時代におけるフレキシブルな字幕制作を実現する研究につなげていきたい。本研究を行うにあたり、聴覚障害者の皆様、日本テレビ放送網株式会社、財団法人NHKエンジニアリングサービス、通信・放送機構渋谷上原リサーチセンターの各位には多大な協力を頂いた。ここに深く感謝する。

〔文 献〕

- 1) 沢村英治, 棚橋大介, 江原暉将, 白井克彦: “実験用字幕つきTVニュースの制作と字幕表示法に関する予備評価”, 映情学誌, 53, 4, pp. 634-638 (1999)
- 2) 門馬隆雄, 沢村英治, 江原暉将, 白井克彦: “多様な提示方法の実験字幕番組の自動制作手法と字幕評価実験概要”, 映情学技報, MIP '99-56, AIT '99-70, 23, 33, pp. 13-18 (1999)
- 3) 沢村英治, 門馬隆雄, 三橋哲雄, 江原暉将, 白井克彦: “多様な提示法の字幕番組評価実験結果と実験用字幕番組の制作概要”, 信学技報, HCS 99-45, 99, 384, pp. 31-38 (1999)
- 4) 沢村英治, 門馬隆雄, 福島孝博, 丸山一郎, 江原暉将, 白井克彦: “聴覚障害者のための字幕放送制作技術の研究状況(第三報)”, 映情学学次大, pp. 115-116 (1999)
- 5) 長田昌次郎, 安田稔, 長谷川敬, 吉田辰夫, 福田忠彦: “文字放送の画像の好ましいコントラスト”, テレビ誌, 39, 6, pp. 516-522 (1985)
- 6) 小峰一晃, 星野春男, 磯野春雄, 内田忠之, 岩鼻幸男, 榊原光裕: “内容了解度によるリアルタイムニュース字幕の評価”, 映情学誌, 52, 4 pp. 504-511 (1998)
- 7) 竹内啓, “統計学辞典”, p. 398, 東洋経済 (1989)

もんま たかお
門馬 隆雄 1995年、早稲田大学理工学研究科電気工学専攻修士課程修了。同年、同大学メディアネットワークセンター助手。1998年より、通信・放送機構渋谷上原RC招聘研究員。自動字幕番組制作システム、字幕画面制作、字幕提示方法の研究に従事。正会員。

さわむら えいじ
沢村 英治 1960年、岩手大学電気工学科卒業。同年、NHK入局。放送用映像機器の研究などに従事。1994年、NHK-ESにてハイビジョンAVシステムの設計・施行に従事。1996年より、通信・放送機構渋谷上原RC研究員として、字幕番組制作システム、字幕提示方法の研究に従事。正会員。

みはし てつお
三橋 哲雄 1964年、電気通信大学卒業。同年、NHKに入局。福岡放送局を経て、1967年より放送技術研究所にて、ハイビジョン、立体テレビの方式、システム、機器開発や視覚心理特性、バリアフリー放送技術などの研究開発に従事。現在、尚美学園大学教授。通信・放送機構渋谷上原RC研究フェロー。正会員。

えはら てるまさ
江原 暉将 1967年、早稲田大学第1理工学部電気通信学科卒業。同年、NHK入局。1970年より放送技術研究所勤務。かな漢字変換、放送衛星の管制制御、機械翻訳、音声認識などの研究に従事。通信・放送機構渋谷上原RCサブリーダー。工学博士。正会員。

しらい かつひこ
白井 克彦 1963年、早稲田大学理工学部電気工学科卒業。1968年、同大学理工学研究科博士課程修了。同年、早稲田大学理工学部電気工学科専任講師。1975年、同教授。1991年、情報学科教授。1994年、教務部長。1998年、常任理事・副総長。通信・放送機構渋谷上原RCリーダー。工学博士。