

第 24 回年次大会予稿

## 動画における字幕デザインのパーソナライズ要素の検討 Subtitle design for personalized video clips

稲垣 洸雄<sup>†</sup> 松村 敦<sup>††</sup> 宇陀 則彦<sup>‡</sup>

Hironori INAGAKI Atsushi MATSUMURA Norihiko UDA

<sup>†</sup> 筑波大学大学院 図書館情報メディア研究科

Graduate School of Library, Information and Media Studies, University of Tsukuba

〒 305-8550 茨城県つくば市春日 1-2

E-Mail: s1621598@u.tsukuba.ac.jp

<sup>††</sup> 筑波大学大学院 図書館情報メディア研究科

Graduate School of Library, Information and Media Studies, University of Tsukuba

〒 305-8550 茨城県つくば市春日 1-2

E-Mail: matsumur@slis.tsukuba.ac.jp

<sup>‡</sup> E-Mail: uda@slis.tsukuba.ac.jp

近年の動画制作において、見やすい字幕の構築が重視されてきている。これは、モバイル端末の普及により、音がなくても内容を理解できるようにするためである。この字幕に対し、それぞれの視聴者にとって、最適な速度や文章サイズといった字幕デザインをパーソナライズして提供できれば、より良い動画視聴体験を提供できると考えられる。そこで本研究では、動画の字幕表示におけるパーソナライズが必要な要素を明らかにすることを目的とし、動画視聴実験を行った。本稿ではその結果を報告する。

The easy-to-read subtitles have become increasingly important in recent video production, because the spread of smartphones requires to understand the video contents without voice. Personalized subtitle design, such as the optimum speed and typography for each viewer makes possible to provide a better video viewing experience. In this paper, we report the results of video viewing experiment to clarify the elements that should be personalized in the subtitle of the video.

キーワード: 動画, 字幕, パーソナライズ

Keywords: video, subtitle, personalize

### 1 はじめに

動画は、我々が何らかの情報を入手する際に、頻繁に目にするものの一つである。特に近年は、モバイル端末における通信回線速度の向上により、いつでも・どんな場所でも動画コンテンツを視聴することができるようになった。モバイル端末でどこでも動画を再生できるとはいえ、外で再生する際、特に公共の空間では音を出さない場合が多い。また、主流な SNS (例えば Facebook や Twitter, Instagram) は、動画を無音の状態で自動的に再生する仕様になっている。そのため、動画コンテンツの作成において、音を出せない状況でも内容が理解できるように、字幕が重要視されている。

字幕作成において重要になるのが、動画内におけ

る文章の可読性である。静止している画面に対する文章とは違い、表示時間や表示スペースといった制約がある中で、文章を配置しなければならない。また、Web 記事や新聞、雑誌などは、スクロールやページめくりといった操作を行うことで、自分の好きなタイミングで情報を取得することが可能である。しかし動画は、情報を表示するタイミングを、コンテンツそのものが操作しているために、字幕を表示する際は、人間が文章を読む速度に合わせた動画編集が必要である。

動画に表示される情報の理解は人によって異なり、文章を読む速度や、一度に表示すべき文字量というのは、個人によって差がある。よって動画編集において、字幕のように文章を表示する際には、文章を

読む速度が遅い人、あるいは一度に表示する文字量について、少ないものを好む人に合わせて構成されている。しかし、そのようにして編集された動画は、文章を読む速度が速い人、あるいは一度にたくさんの文章を読むのを好んでいる人にとって望ましいとは言えない。

そこで、視聴者の字幕表示の嗜好に合わせて動画を提供することで、より良い動画視聴体験を実現することが可能と考えられる。

ただ、その際問題となるのが、字幕の具体的ななどの要素を、どのようにして変化させることで、視聴者の字幕表示の嗜好に合わせることができるといことである。なぜなら、字幕の主観的な読みの嗜好について焦点を当てた研究は未だに無く、どのような字幕の要素を視聴者ごとに最適化するべきかわかっていないためである。

字幕の変更できる要素の仮説として考えられるのは、例えば動画の「速度」や、「1行あたりの文字数」「行数」といった字幕のデザイン要素がある。これらの要素のパラメータについて、多くの人が同じパラメータのものを見やすいと判断するか、あるいはそれぞれの人が別々のパラメータの字幕を見やすいと判断するかを明らかにすることで、パーソナライズが必要か否かを判断できると考えた。

そこで本研究では、動画の字幕表示におけるパーソナライズが必要な要素を明らかにすることを目的とした。

本論文は次のように構成されている。第2章では先行する研究を述べる。第3章では本研究で用いた実験手法について述べる。第4章では実験結果に対する考察を述べる。最後に、第5章では本論文の結論を述べる。

## 2 関連研究

石井らの研究<sup>[1]</sup>では、携帯端末などのデザインに資することを目的として、横書きの文章が下から上へと流れる縦スクロール表示について、その基本的な読み特性を調査している。この研究では、文章における読み特性の評価指標として、単位時間あたりの文章の読み速度を、文字単位で計算したものを快適速度[文字/秒]として利用している。

研究結果として、「移動単位が1ピクセルの条件の方が快適速度が速いこと、また表示行数および表示文字数の増加にともない快適速度が速くなること、

表示枠内の広さが同じであれば、行数を犠牲にして行内の文字数を多くとる方が快適速度が速くなること」が示されている。また、「読み最中の眼球運動を測定した結果、移動単位により読み方が異なること、また、移動単位によらず読み手は表示枠内の最下行に視線を向けて読むことが多いこと」が明らかになったとされている。

向後らの研究<sup>[2]</sup>では、映像と文字情報の提示についてのデザイン原則への示唆を得ることを目的として、外国映画の字幕つきビデオを視聴するときの眼球運動を分析している。その結果、「字数が増えれば増えるほど文字あたりにかかる時間が長くなり、そこでは行数の要因と文字数の要因とが効いていること」が明らかにされたとしている。

これらの研究で調査されたのは、人間の基本動作や、デザイン原則を明らかにするために、読み時間の速度に関する情報を眼球運動の測定といった手法によって明らかにしようとしたものである。

これに対して、本研究は、動画内文章の快適な読み実現のためのパーソナライズを目標としているため、動画内の文章に対する、動画視聴者の主観的な見やすさというものに着目している。

## 3 実験

### 3.1 実験目的

本実験は、動画字幕におけるパーソナライズが必要な要素を明らかにすることを目的とする。

### 3.2 実験参加者

18歳から25歳までの大学生及び大学院生19人が実験に参加した。実験参加者は全員、裸眼または矯正で正常な視力であった。

### 3.3 環境

実験に使用した動画再生機器はMacbook Pro 13-inch with Retina Displayを使用した。実験参加者は、実験開始時にディスプレイから顔までの観察距離が70cmとなる位置に座り、顔面の固定などを行わず、実験を行った。



### 3.4 実験システム

本研究では、動画内文章の主観的見やすさを検証するために、実験参加者が利用する、速度調節が可能な動画再生ツールを開発した。また、多くの動画を手動で作るには手間がかかるため、文章を入力することで字幕動画を自動的に生成するシステムを開発し、出力された動画を実験に使用した。

### 3.5 実験参加者への刺激

実験に使用したディスプレイの画面中央付近に表示枠を設け、そこで動画を再生した。動画の内容は、Web上の動画サービスで多く採用されている字幕表示方式で表現される字幕を再現した動画である。

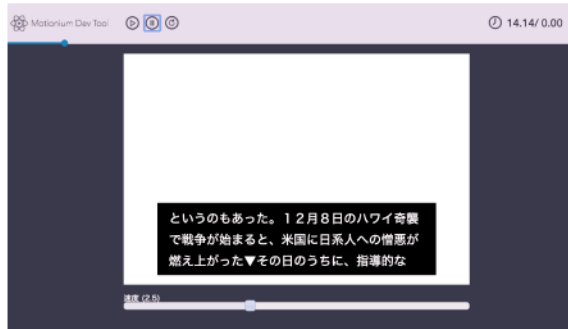


図 1: 実験システムのスクリーンショット

図1が実験システムのスクリーンショットである。字幕は画面下部に黒い帯を表示し、そこに白色で文章を表示する。動画は、この字幕が一定の時間間隔で切り替わるようになっている映像である。

ディスプレイのモニタ表示領域は横 286mm、縦 179mm、画面解像度は 2,560px × 1,600px であった。字幕に表示する文章の 1 文字の大きさは 26px(縦横 6mm) であった。

文字のフォントはゴシック体とした。行間は 1mm 空けた。

字幕の表示条件として、1 行の表示文字数は、10, 15, 20 文字の 3 条件、1 つの字幕内に表示する表示行数は 1, 2, 3 行の 3 条件であった。

### 3.6 文章材料

朝日新聞社編・天声人語 (2015 年) の中から、時事内容のものを避けて選定したものを、練習試行用に 1 文章、本番試行用に 9 文章用いた。

### 3.7 手続き

実験参加者の課題は、動画内の字幕を読みながら、文章を読むのに対して最も読みやすいと感じる動画速度に調節することである。実験参加者には、動画が終わる前に速度の調整を行い、動画終了時の速度を記録することを教示した。

また、実験開始の前に、読んだ文章についての簡単な確認テストを行うことを教示した。どの動画についての確認テストを行うかについては事前に知らせずに、どの試行で確認テストが行われても回答できるように、字幕文章の内容を理解しながら動画を視聴するように求めた。確認テストは、全試行終了後に、確認を行う動画で出現したキーワードを幾つか例示し、動画の内容について説明してもらうものである。

1 試行の流れは以下のとおりである。まず、図 1 に示すように、画面上には動画の表示枠と再生コントローラ、速度調整用のバーが配置されている。実験参加者は再生コントローラの再生ボタンをクリックすることで、動画の再生が始まる。動画の内容は、字幕が表示されたページが一定時間ごとに切り替わるというものである。開始時のページ切り替え時間は 10sec/page に設定されている。次に、ページ切り替え時間の速度調整用バーをクリックする。実験参加者は、動画内の字幕文章を読みながら、キーボードの右キー (加速) と左キー (減速) を用いて自身が読みやすいと感じる速度にまで調整する。動画が最後まで流れることで 1 試行が終了とし、その時点での動画速度を記録する。

速度調整時にキーボードを一回押した時の変化量は、初期速度から 0.1 倍ずつ増加、もしくは減少する。例えば、ページ切り替え時間の初期値である 10sec/page から 1 回右キーを押すと、速度は 1.1 倍となり、ページ切り替え時間は 9.1sec/page となる。その後、9 回右キーを押すことで、倍率が 2 倍になり、5sec/page となる。

### 3.8 デザイン

本実験は、デザインの変更要素として行数と 1 行の表示文字数の変更を行った。表示する行数として 3 水準 (1, 2, 3 行)、1 行の表示文字数として 3 水準 (10, 15, 20 文字) の 9 条件からなる動画を生成し、実験に使用した。

また、長い文章を字幕化するにあたり、1字幕ごとに文字を切り分ける必要がある。そこで今回は、表示最大文字数から3文字以内に出現した、単語の意味的な切れ目、もしくは助詞の直後、句読点という要素において、文章を区切る処理を行った。

最初に練習試行を1試行、その後本試行を9試行行った。練習試行では、実験で用いる動画再生ツールの操作方法に慣れることを目的として、本試行では表示していない動画で行ってもらった。

### 3.9 インタビュー調査

実験参加者の属性を調査することで、パーソナライズに有効な属性と動画内要素を検証するために、実験後にインタビュー調査を行った。インタビューでの質問は大きく分けて4つの質問で構成されている。内容は、実験参加者自ら速度調節した後の動画で気になった点について、動画に対しての要望について、普段の動画の視聴習慣について、普段どの程度文章を読んでいるかについてを問うものである。

## 4 実験結果と考察

実験の結果より、動画字幕におけるデザインと、主観的な読みやすさの評価を比較することにより、動画中における文章要素のパーソナライズの必要性について考察する。

### 4.1 快適速度と主観的な読みやすさの関係

実験参加者19人分の動画の切り替え時間と主観的な見やすさの実験結果から、それぞれの関係を表したものを図2のグラフに示した。このグラフでは、快適速度における単位秒あたりの文字数を横軸に、主観的な見やすさの5段階評価が縦軸に取られている。横軸に関しては、数字が大きくなるほど字幕を読む速度が速かったということになる。

多くの実験参加者のグラフにおいて、ほとんど同じ快適速度における単位秒あたりの文字数においても、評価が高いものと低いものが混在していることがわかる。また、主観的な見やすさとの関係において、何らかの相関性を持っているようには分析できない。

### 4.2 字幕デザインと主観的な読みやすさの関係

字幕デザインと主観的な読みやすさの関係を考察するために、字幕の1行の横幅と行数の2つのデザイン的要素に対する、主観的な読みやすさの評価が最も高かった実験参加者のIDをまとめた表を以下に示す。

表1: 最も読みやすかった字幕デザインのグループ

行数	1行あたりの文字数				
	10	15	20	10-15	15-20
1	19	17	2,8	-	-
2	3,12,18	4,9,14	-	16	15
3	-	11	5,13	-	-
1-3	-	-	7	-	-
2-3	1,10	6	-	-	-

行数における主観的な読みやすさについて比較すると、行数が2行の時に最も見やすかったという結果が合計で8人おり、ある程度集中していると言える。しかしながら、1行の時に最も見やすいという結果の人が4人、3行の時に最も見やすかったという結果の人が3人おり、極端に2行のものが良かったという偏りがあるとは言えない。

同様に、1行あたりの文字数における主観的な読みやすさについての比較でも、最も見やすかったという結果が出た人数について、文字数が10文字の時は6人、15文字の時は6人、20文字の時は5人と、こちらも分散した結果となった。

すなわち、1行あたりの文字数と行数といった字幕のデザイン要素においては、どの視聴者も一番見やすいというものは存在せず、個人の嗜好によって大きく左右される要素であると考えられる。

### 4.3 実験参加者の属性と主観的な見やすさの関係

本研究では、動画中における文章の快適な読みを実現するためのパーソナライズ要素の検討を掲げた。パーソナライズの実現には、視聴者から何らかの情報を受け取り、それを最適な動画の生成に生かす必要がある。そこで、日頃からどの程度テレビやインターネット動画、またはニュース番組を視聴するか、

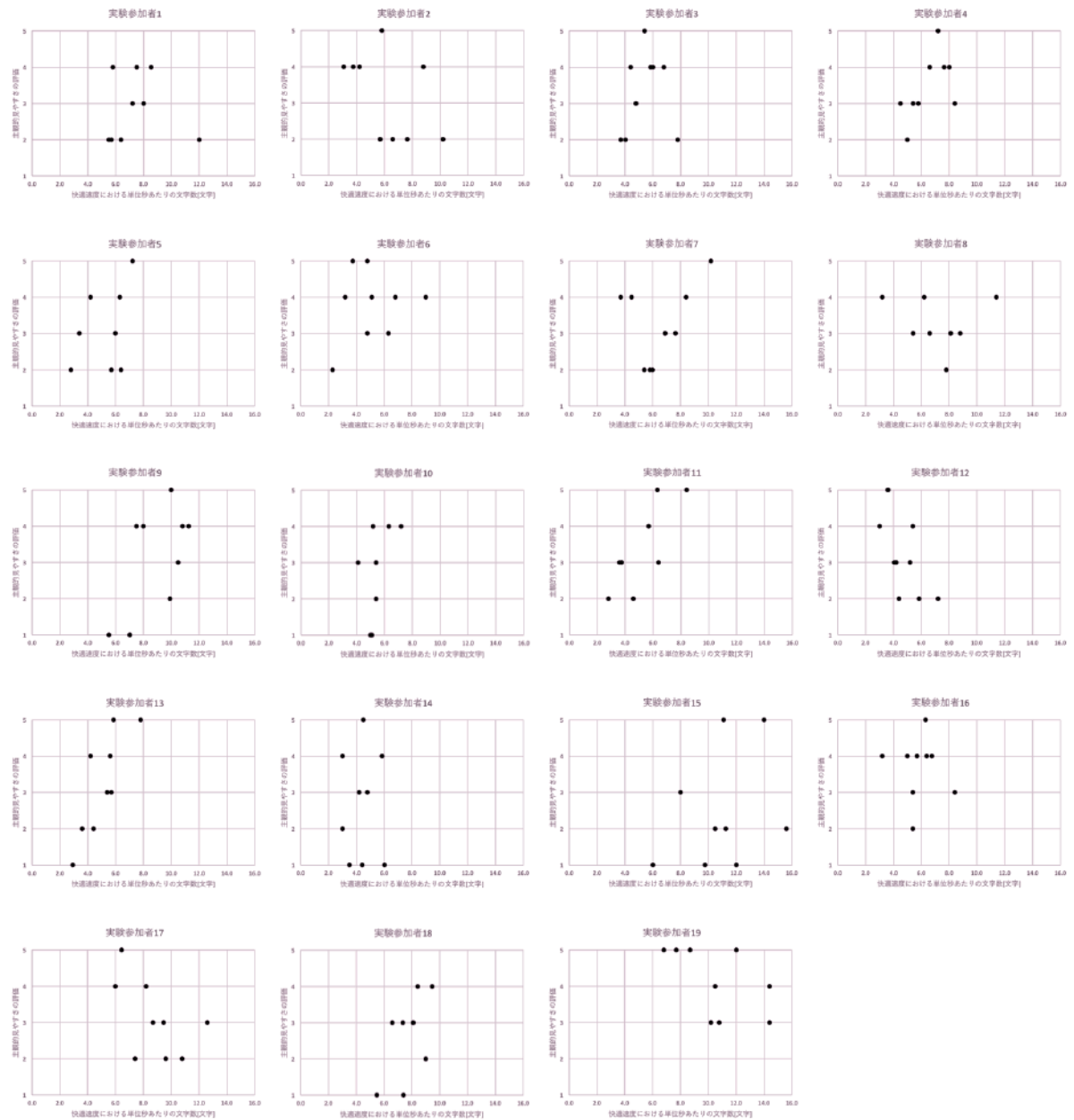


図 2: 動画の文字を読む快適速度と主観的見やすさの関係



あるいは、日頃からどの程度読書をしているか、新聞を読んでいるか、インターネット上の記事を読んでいるかなどの情報を収集した。しかし、読み取れる実験参加者の属性からは、主観的な読みやすさとなんらかの相関関係にある項目を見つけることができなかった。

#### 4.4 今後の課題

今回の実験では、実験参加者の属性と主観的な読みやすさについて明らかにすることができなかったために、具体的に動画内文章のパーソナライズを実現するための手法まで提案することはできなかった。そこで今回調査した、日頃の動画の視聴習慣、及び日頃から読む文章の量について以外の点でも、実験参加者の属性調査を行うことで、主観的な読みやすさがどんな属性に依存するのかを明らかにする必要がある。

また、字幕を構成する文章作成についても、実験参加者が見やすいと感じる文章を構成する必要があると考えられる。

今後は、実際にパーソナライズを行った際に、利用者の満足度にどの程度の効果があるのかの調査を行う必要がある。

今後の課題として、利用者の字幕デザインの嗜好と利用者の属性との間にある関係を明らかにすることが挙げられる。また、実際にパーソナライズを行うシステムを構築し評価を行う必要がある。

## 参考文献

- [1] 石井亮登, 森田ひろみ. “縦スクロール表示された文章の快適な読み速度と眼球運動”. 情報処理学会論文誌, 2013. Vol.54, No.6, p.1784-1793.
- [2] 向後千春, 岸学. “字幕映画の視聴における眼球運動の分析”. 日本教育工学雑誌, 1996. Vol.20, No.3, p.161-166.

## 5 結論

本研究では、動画における字幕デザインのパーソナライズ要素を明らかにするために、複数の字幕デザインを提示する動画視聴実験を行った。実験参加者には、動画内の字幕の行数と1行あたりの文字数が異なる9つの動画を、自身が見やすいと思う速度に調整しながら見てもらった。

実験の結果、パラメータを操作して作成した複数の字幕デザインに対して、どの視聴者も一番見やすいというものは存在せず、字幕デザインの嗜好は個人ごとに差があることが明らかになった。このことは、行数と1行あたりの文字数という字幕デザインの要素をパーソナライズすることの有効性を示していると考えられる。一方で、実験で示された動画字幕の切り替え速度と主観的な見やすさの関係から、字幕を読む速度と主観的な読みやすさに相関はなかった。すなわち、主観的な読みやすさは、読み速度よりも字幕デザインによって影響される可能性が示されたと言える。