ARグラスを用いたろう・難聴者向け字幕提示方法の探索的研究：字幕位置の影響に着目して

船山滉介[1,a)]　設楽明寿[1,2,b)]　米山文雄[1,c)]　加藤伸子[1,d)]　白石優旗[1,e)]

概要：ろう・難聴者向けの情報保障手段として，AR グラスが注目されつつある．従来の HMD 技術では，単に字幕をミラーリングしており頭の動きに合わせて対象物が追従されている．一方，AR グラスは，従来の HMD 技術だけではなく字幕の提示位置を AR 空間上でユーザーが柔軟に設定できる技術も持ち合わせている．本研究の目的は，字幕の表示位置の影響に着目し，1）従来の HMD 技術（ミラーリング），2）AR 空間上で表示位置を固定，3）AR 空間上でユーザーが表示位置を自由に決定，の 3 種類の字幕提示方法で実験し，映像コンテンツと対人コミュニケーションの内容理解度に与える影響を明らかにすることである．また，各手法の課題と利点について実験結果とフィードバックに基づいて詳細に分析することで，AR グラスを用いた字幕提示方法の可能性と改善点を探ることを目指した．実験方法は，e-sports の試合映像を視聴しながら，隣の手話話者とコミュニケーションをとる環境下で，それぞれの字幕提示方法を試した．その後にそれぞれの字幕提示方法での内容理解度を評価し，それぞれの手法の課題と利点についてもフィードバックで収集した．

キーワード：AR，AR glass，e-sports，情報保障，ろう・難聴者

An Exploratory Study on Caption Display for Deaf and Hard of Hearing Users with AR Glasses: Effects of Caption Positioning

Funayama Kosuke[1,a)]　Shitara Akihisa[1,2,b)]　Yoneyama Fumio[1,c)]　Kato Nobuko[1,d)]
Shiraishi Yuhki[1,e)]

Abstract: AR glasses offer new possibilities for caption accessibility for deaf and hard of hearing individuals. Unlike conventional HMDs that mirror captions, AR glasses allow flexible positioning of captions in AR space. This study examines the impact of caption positioning on video content comprehension and interpersonal communication. We compare three methods: 1） conventional HMD mirroring, 2） fixed AR positioning, and 3） user-determined AR positioning. Our aim is to explore the potential and limitations of each method through experimental analysis and user feedback. Participants watched e-sports footage while communicating with a sign language speaker, using each caption method. We then assessed content comprehension and gathered feedback on the advantages and challenges of each approach.

Keywords: AR, AR glass, e-sports, information accessibility, deaf and hard of hearing

1. はじめに

ろう者や難聴者のための情報保障の 1 つとして，HMD（ヘッドマウントディスプレイ）を用いられる．従来の

1 筑波技術大学
Tsukuba University of Technology, 4-3-15 Amakubo, Tsukuba, Ibaraki 305–8520, Japan
2 筑波大学
University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8577, Japan
a) a211130@a.tsukuba-tech.ac.jp
b) theta-akihisa@digitalnature.slis.tsukuba.ac.jp
c) yonefumi@a.tsukuba-tech.ac.jp
d) nobuko@a.tsukuba-tech.ac.jp
e) yuhkis@a.tsukuba-tech.ac.jp

© 2024 Information Processing Society of Japan

HMD 技術では，単に字幕をミラーリングしており，頭の動きに合わせて対象物が追従されている．一方，AR グラスは，従来の HMD 技術だけではなく，字幕の提示位置を AR 空間上でユーザーが柔軟に設定できる技術も持ち合わせており，様々な方法で字幕提示が可能である．

本研究では，字幕の表示位置の影響に着目し，以下の 3 種類の字幕提示方法で実験を行った．

1： 従来の HMD 技術（ミラーリング）：単に字幕をミラーリングし，頭の動きに合わせて対象物が追従する．
2： AR 空間上で表示位置を固定：字幕の位置を AR 空間上で固定して表示する．
3： AR 空間上でユーザーが表示位置を自由に決定：字幕の位置を AR 空間上でユーザーが自由に変更できる．

これらの字幕提示方法を用いて，動画コンテンツと対人コミュニケーションの内容理解度に与える影響を明らかにすることを目的とした．

本研究では，実験に用いる AR グラスは XREAL Air[1] を選定した．その主な理由は，XREAL 社が公開されている NRSDK1.9.0（開発当時のバージョン）を活用した Unity を用いた開発が可能だからである．AR モードで動作するアプリを制作することで，3 種類の字幕提示方法に基づいたシステムを開発した．

また，各字幕提示方法の課題と利点について実験結果とフィードバックに基づいて詳細に分析することで，AR グラスを用いた字幕提示方法の課題と利点を探ることを目指した．調査項目として，以下の 3 つのリサーチクエスチョン（RQ）を設定した．

RQ1. 動画コンテンツの内容を把握することができるか
字幕提示方法が動画コンテンツの内容理解に与える影響を調査する．

RQ2. 対人コミュニケーション（手話）の内容を把握することができるか
字幕の表示位置や表示方法が，手話の視認の妨げにならないかを調査する．

RQ3. 各字幕提示方法の利点と課題はなにか
ユーザーの体験後のフィードバックアンケートと実験全体を通してのフィードバックアンケートから, 各字幕提示方法の特徴や改善点を調査する．

2. 関連研究

HMD を活用した研究において，西岡ら [2] は講義の情報保障における PC による字幕提示と HMD の比較から，HMD が視線移動の負担を軽減する効果があることを確認した．さらに，加藤ら [3] は博物館の鑑賞を模擬した実験で，対象と字幕の同時視認における課題を指摘している．また，阪田ら [4] は AR 空間上にテキストを提示しジャイロセンサを用いることで別の動作を行うことができ，円滑なコミュニケーションが可能であると述べている．

国外の研究では，Vinoba Vinayagamoorthy ら [5] は，AR 上での手話通訳の提供がろう・難聴者にとってのテレビニュース視聴における情報保障のレベルを向上させる可能性があると報告している．また，AR 上で提示される情報源の位置を個々人が求める位置に調整できるというパーソナライゼーションの重要性も示唆されている．

Larwan ら [6] は，参加者の字幕スタイルの好みについて，「研究結果は，伝統的な『TV CC Style』と『Movie Subtitle Style』の字幕に対する参加者の好みがほぼ二分していたことを明らかにした」と述べている．この結果は，聴覚障害者の字幕スタイルの好みが一様ではなく，個人差が大きいことを示唆している．また，Larwan ら [6] は，「参加者は，字幕をビデオの内側（下側）に表示することを強く好んだ」ことも報告している．この結果は，字幕の位置が視聴体験に大きな影響を与える可能性を示唆している．しかし，この結果は Crabb ら [7] の研究とは異なるものであった．オンラインビデオの字幕はビデオフレームの外側に配置することを推奨していた．

本研究では，ミラーリングだけではなく，AR 空間を活用した字幕提示方法をアプリとして制作し，各字幕提示方法の利点や課題を当事者視点も含めて調査する．また，各字幕提示方法による対人コミュニケーションの影響についても明らかにすることを目指す．

3. 実験方法

3 種類の字幕提示方法と 3 種類の試合動画・手話動画を用いて比較実験する．実験参加者は，これらの字幕提示方法を全て体験（計 3 回）行い，テスト・フィードバックをしてもらう．また，実験の様子をカメラ・ビデオカメラで写真・動画として記録する．字幕提示の種類（小文字英字の a, b, c）は以下の通りに定義する．

実験で使用した字幕提示方法は以下の 3 種類である：

a ：従来の HMD 技術（ミラーリング）
b ：AR 空間上で表示位置を固定
c ：AR 空間上でユーザーが表示位置を自由に決定

3.1 実験環境

各字幕提示方法を評価するために，e-sports の試合を観戦しながら隣の人と会話をしている環境を再現した．この環境は，隣に座る友人や家族などとペアでスポーツ観戦を行ったときの対人コミュニケーションが発生する状況から着想したものである．

e-sports では，注視する対象（モニター）が移動せず，観戦者も特別な事情がない限り試合中は移動することがないことが想定される．また，観戦者同士でコミュニケーションをとる機会が多いため，対人コミュニケーションの状況

表 1　試合動画，手話動画の長さ

	試合動画	手話動画（休憩 1 回目）	手話動画（休憩 2 回目）
A	6 分 43 秒	30 秒	50 秒
B	7 分 36 秒	24 秒	34 秒
C	7 分 20 秒	33 秒	38 秒

各試合動画のそれぞれに対応する手話動画を提示する．

を再現するのに適している．会場の環境は薄暗いが，手話でコミュニケーションをとるのに十分な明るさが確保されている．そのため，3 つの RQ を評価することができる．

実験参加者は，AR グラスを装着し，e-sports の試合を観戦しながら，隣の手話話者と会話を行う．これにより，AR グラスを用いた字幕提示方法が，対人コミュニケーションにどのような影響を与えるかを分析することができる．

本研究では，AR グラスを用いた字幕提示方法を評価・分析するために，図 1 のように e-sports 観戦を想定した実験環境を構築した．

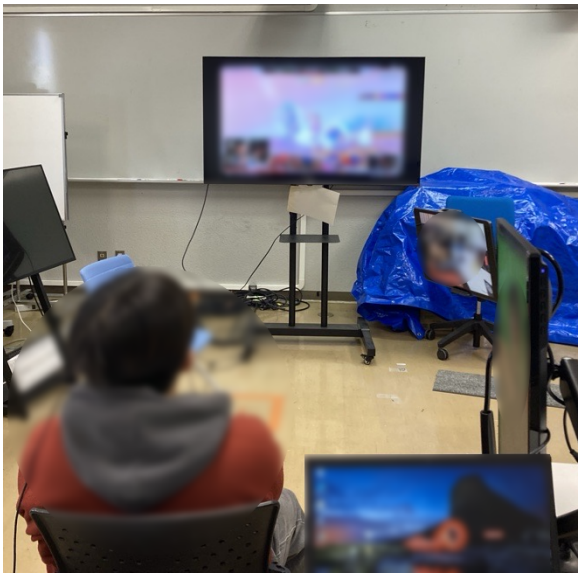

図 1　e-sports 観戦を想定した実験環境の様子

e-sports の試合動画とそれに対応する字幕を用意した．字幕は，srt ファイルでタイムスタンプを作成し，動画開始時に実験実施者がボタンを押すことで動画と同期させた．字幕の内容は，誤字・脱字のない，実況・解説者の音声を文字起こししたものである．

隣の手話話者とのコミュニケーションを再現するために，実験参加者の隣に手話動画を提示するモニターを設置した．モニターと実験参加者の距離は 30cm とした．手話動画は，試合動画内の 1 ラウンド終了時と 2 ラウンド終了時のインターバルに合わせて，モニター提示されるようにした．そして，実験実施者が実験参加者の肩をたたいてモニターを見るように誘導した．

3.2　実験に用いた動画

実験で使用した e-sports 試合動画，各試合動画のそれぞれに対応する手話動画の長さを表 1 に示す．

以下に毎実験のタイムラインを提示する．

図 2　試合動画，字幕，手話動画のタイムライン

図 2 は，1 回の実験における試合動画，字幕，手話映像の提示タイミングを示している．試合中は試合動画（黄色）が表示され，休憩時には会場の画面と手話動画（緑色）が表示される．字幕（青色）は常時表示されている．

3.3　実験手順

実験の手順は以下の通りである．

1： 事前理解テストを最初に実施する．次に，補聴器または人工内耳を装用している参加者は，これらを外すか電源を切る．その後，「XREAL Air」を装着する．

2： PowerPoint を用いて被験者に実験用アプリの操作説明を行い，練習する．

3： 字幕提示方法（a, b, c）と試合動画（A, B, C）の組み合わせに基づいて，試合動画と手話動画を含む 10 分間視聴を行う．

4： 評価の 15 分では，試合動画内容理解度テスト，手話動画内容理解度テスト，各字幕提示方法のフィードバックを行う．これを 3 回行う．

5： 全体の実験が終了した後，実験全体のフィードバックを行う．

順序による影響を防ぐためにラテン方格法で実施した．また，実験設計に基づいて字幕提示方法（a，b，c）の組み合わせを網羅かつ，順位効果を軽減させるために必要最小限の人数は 12 名とした．試合動画の順序効果は考慮されていない．試合動画と字幕提示方法の組み合わせは，表 2 に示す．なお，表 2 の A, B, C はそれぞれ試合動画 A，B，C を表し，a, b, c はそれぞれ字幕提示方法 a, b, c を表している．

3.4　実験参加者

Microsoft Forms を活用し，本学学生 20 代のろう・難聴者 12 名に実験協力の公募を行い，12 名の協力が得られた．参加者の属性について性別は男性 12 名，女性 0 名であった．また，補聴器/人工内耳の装用の有無について「補聴器のみ装用」が 6 名，「人工内耳のみ装用」が 5 名，「どちらも装用しない」が 1 名であった．なお，実験参加者の条件は「裸耳で概ね 60dB 以上」「日常生活で手話を用いたコミュニケーションが可能である」「実験時に裸眼または，コンタクトレンズを装着可能」とした．眼鏡装用であれば AR グラス装用時に歪みやピントが合わせづらく実験環境

© 2024 Information Processing Society of Japan

表 2　試合動画と字幕提示方法の組み合わせ

ID	1 回目	2 回目	3 回目
1	Ba	Cc	Ab
2	Aa	Bb	Cc
3	Cb	Aa	Bc
4	Ab	Cc	Ba
5	Bc	Aa	Cb
6	Cc	Bb	Aa
7	Ca	Ab	Bc
8	Ca	Bc	Ab
9	Bb	Ca	Ac
10	Bb	Ac	Ca
11	Ac	Ba	Cb
12	Ac	Cb	Ba

を均一することが難しくなるため，含まなかった．

3.5　解析方法

本実験では，以下の 4 種類のデータを収集した：

1： 試合内容理解度テスト（A, B, C）
2： 手話内容理解度テスト（A, B, C）
3： 各字幕提示方法体験直後のアンケート
4： 実験全体終了時のアンケート

テストの採点方法については，すべての問題を 1 問 1 点として評価した．試合内容理解度テストは合計 13 問で構成され，最高得点は 13 点である．一方，手話内容理解度テストは合計 6 問で構成され，最高得点は 6 点である．データ解析においては，主に各字幕提示方法間の比較分析を行った．

具体的には，各字幕提示方法に対応するデータセットを個別に抽出し，それぞれの方法間で統計的比較を実施した．さらに，アンケート結果の分析を通じて，各字幕提示方法に対する参加者の主観的評価や全体的な実験体験に関する定性的データを収集した．

4.　実験結果

本実験では，試合内容理解度テストのスコア．手話動画内容理解度テストのスコア．各字幕提示方法体験直後のアンケート回答データ．および実験全体終了時のアンケート回答データの 4 種類のデータを収集した．なお，以降の表や図のみ，視認性を考慮して字幕提示方法 a, b, c を提示方法 a, b, c と省略して記載している．

4.1　試合内容理解度テスト解析

試合動画 A, B, C における試合内容理解度テストの結果を分析するため，3 つの試合動画の全ての組み合わせ（A-B, A-C, B-C）に対して対応のある t 検定を実施した．表 3 から，いずれの組み合わせにおいても統計的に有意な差は認められなかった (全ての比較で Bonferroni 補正後の p >0.016 が確認された)．

表 3　試合動画別の内容理解度テスト結果

試合動画	M	SD	比較対象	p 値
A	6.42	2.02	B	0.7090
B	6.17	2.12	C	0.6887
C	6.33	2.42	A	0.8742

(M: 平均, SD: 標準偏差, *p <0.05, **Bonferroni 補正 p <0.016)

表 4　字幕提示方法別の試合内容理解度テスト結果

カテゴリ (n)	方法 a		方法 b		方法 c	
	M	SD	M	SD	M	SD
字幕 (4)	1.50	1.00	1.75	0.97	1.75	1.42
休憩中の字幕 (1)	0.42	0.51	0.58	0.51	0.33	0.49
映像 (4)	1.25	1.29	1.50	1.09	1.08	0.79
字幕と映像 (4)	2.92	1.24	2.83	1.11	3.00	0.95

((n): カテゴリ別の問題数,M: 平均, SD: 標準偏差)

表 5　手話動画別の内容理解度テスト結果

手話動画	M	SD	比較対象	p 値
A	4.25	0.62	B	0.0105**
B	5.08	1.00	C	0.4569
C	4.75	0.97	A	0.2143

(M: 平均, SD: 標準偏差, *p <0.05, **Bonferroni 補正 p <0.016)

4.1.1　字幕提示方法間の試合内容理解度比較

試合内容理解度テストの結果を字幕，手話動画視聴時に流れていた字幕（休憩中の字幕），映像，および字幕と映像の組み合わせの 4 つのカテゴリに分けて分析した．各カテゴリにおける 3 つの字幕提示方法（a， b， c）の平均スコアと標準偏差を表 4 に示す．

全てのカテゴリにおいて，3 つの字幕提示方法間で統計的に有意な差は認められなかった．

特に，手話動画視聴時に流れていた字幕カテゴリについては，二項検定を用いて正解数の比較も行った．方法別の正解数は，方法 a: 5/12，方法 b: 7/12，方法 c: 4/12 であったが，方法間で有意な差は見られなかった．

4.2　手話内容理解度テスト解析

手話動画 A，B，C における手話内容理解度テストの結果を分析した．各手話動画の平均スコアと標準偏差，および対応のある t 検定の結果を表 5 に示す．対応ある t 検定の結果, 手話動画 A と手話動画 B の間に統計的に有意な差が認められた (Bonferroni 補正後の p <0.016).

4.2.1　字幕内容正誤による手話内容理解度の差異分析

手話動画時に流れていた字幕の内容を 3 つの選択肢から正しい選択肢と選ぶことができたかどうかによって，参加者を字幕提示方法ごとに分けて字幕内容正答グループと字幕内容誤答グループに分類し，各字幕提示方法における手話内容理解度を分析した．グループ間の比較には対応なしの t 検定を用いた．結果を表 6 に示す．

対応なしの t 検定の結果，字幕提示方法 a において，字幕内容誤答グループは字幕内容正答グループよりも有意に

© 2024 Information Processing Society of Japan

表 6　字幕内容正誤による手話内容理解度テスト結果

提示方法	グループ	n	M	SD	p 値
a	字幕内容正答	5	4.40	0.89	0.0311*
	字幕内容誤答	7	5.43	0.53	
b	字幕内容正答	7	4.71	1.11	0.6140
	字幕内容誤答	5	4.40	0.89	
c	字幕内容正答	4	4.50	1.00	1.0000
	字幕内容誤答	8	4.50	0.93	

（n: 人数, M: 平均, SD: 標準偏差, *p <0.05）

高い手話内容理解度を示した．一方，字幕提示方法 b および c では，両グループ間に有意な差は見られなかった．

4.3　各字幕提示方法体験直後のアンケート解析

字幕提示方法体験直後，実験参加者に対して 7 段階リッカート尺度および自由記述式のアンケートを実施した．設問数は 7 段階リッカート尺度による評価が 12 問，自由記述式が 1 問の計 13 問である．

4.3.1　字幕提示方法別リッカート尺度（7 段階）の解析

各字幕提示方法（a， b， c）に対するユーザーフィードバックを 7 段階のリッカート尺度（1: 全くあてはまらない ～ 7: 非常にあてはまる）を用いて収集し，対応ありの t 検定を行った．

Q.1 試合の内容を把握することができた．

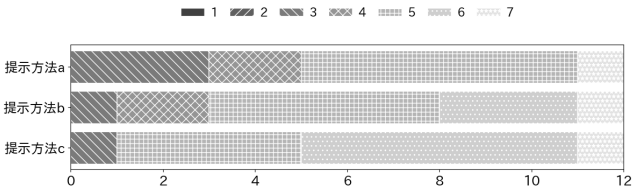

図 3　試合内容の把握に関する回答分布

表 7　試合内容の把握

提示方法	M	SD	比較対象	p 値
a	4.50	1.17	b	0.1315
b	5.08	1.08	c	0.2691
c	5.50	1.00	a	0.0204*

(M: 平均，SD: 標準偏差, *p <0.05, **Bonferroni 補正 p <0.016)

字幕提示方法 c は字幕提示方法 a と比較して，より高い評価を得る傾向が見られた（p <0.05，Bonferroni 補正前）．Bonferroni 補正後（p <0.016）では統計的有意差は確認されなかったものの，有意傾向（Bonferroni 補正後の p <0.033）が観察された．さらに，効果量を算出したところ，大きな効果が確認された（Cohen's d = 0.92, 95% CI [0.19, 1.81]）．回答分布については，図 3 を参照．

4.3.2　自由記述形式のフィードバック解析

さらに，各字幕提示方法に対する自由記述形式のフィードバックを収集し，分析を行った．

字幕提示方法 a（従来の HMD 技術（ミラーリング））

- 字幕が視界の中心に固定されることへの不満が見られた．「視線真ん中に字幕が映り続けるので鬱陶しい」という意見があった．
- 一方で，「TV の位置と字幕の位置が同じだと，ピントを合わせやすく，見やすかったし目が疲れにくかった」という肯定的な意見もあった．

字幕提示方法 b（AR 空間上で表示位置を固定）

- 字幕と手話での情報提示が明確に区別されていたことが評価された．「前回（字幕提示方法 a）とは違って明確に字幕と手話での情報を区別して提示していたので，分かりやすかった」というコメントがあった．
- 字幕位置のカスタマイズ性への要望が依然として高かった．「字幕を好きな位置に配置できるのが良いと感じた．個人差によっては見やすい位置が変わりそう」という意見が多く見られた．

字幕提示方法 c（AR 空間上で表示位置を自由に決定）

- 多くの参加者が字幕位置を自由に設定できることを高く評価した．「字幕の位置を好きに設定できるのはとても良い」，「自分で決めた字幕の位置の方が見やすいと感じた」といったコメントが多く見られた．
- 字幕と映像の同時視聴が容易になったという意見があった．「字幕を映像の右下に置くことで，左目は映像で右目は字幕をチラ見できた」というコメントはこの点を示唆している．
- 一部の参加者は，字幕の奥行き（3D 空間での位置）調整も有用だと感じていた．「字幕の場所を自分で決めることの他に文字を奥にできるという面ではとても見やすかった」という意見があった．
- 微調整の必要性を指摘する声もあった．「ちょっとした微調整はまだ要るようだったが，それでも字幕が固定されている分，ストレスを感じなかった」というコメントがあった．

4.4　実験全体終了時のアンケート解析

実験全体が終了した後，参加者から最終的なフィードバックを収集するためのアンケートを実施した．このアンケートは，字幕提示方法の総合的な評価と将来的な改善点を把握することを目的としている．

4.4.1　字幕提示方法の選択

Q.1 もし，情報保障を受けるのであれば 3 つのうちどちらがほしいですか？

参加者に最も好ましいと感じた字幕提示方法を選択してもらった結果を図 4 に示す．字幕提示方法 c を選択した参加者が 10 人，字幕提示方法 b を選択した参加者が 2 人，字幕提示方法 a を選択した参加者は 0 人であった．

4.4.2　字幕と手話の同時把握

字幕提示方法 a と AR を活用した字幕提示方法 b，c とのでの手話と字幕の同時把握の可能性について，参加者の回答を比較した結果を表 8 に示す．

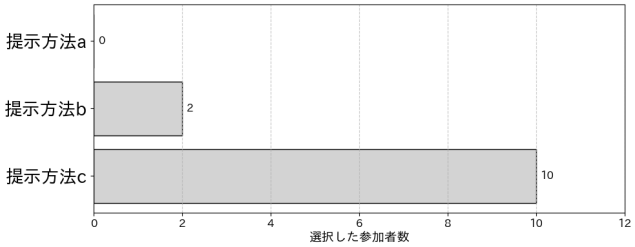

図 4　参加者が選択した字幕提示方法

表 8　手話と字幕の同時受容の可能性

	提示方法 a	AR 版 (提示方法 b, 提示方法 c)
可能	5 人	2 人
不可能	7 人	10 人

4.4.3　字幕と手話へのそれぞれの集中度

参加者に対し，「字幕と手話それぞれの集中度を合計 100 で教えてください」という質問を行った．この集中度は参加者の主観的な評価に基づくものである．

4.4.3.1　字幕提示方法 a での集中度

字幕提示手法 a における字幕と手話への集中度の分布を図 5 に示す．

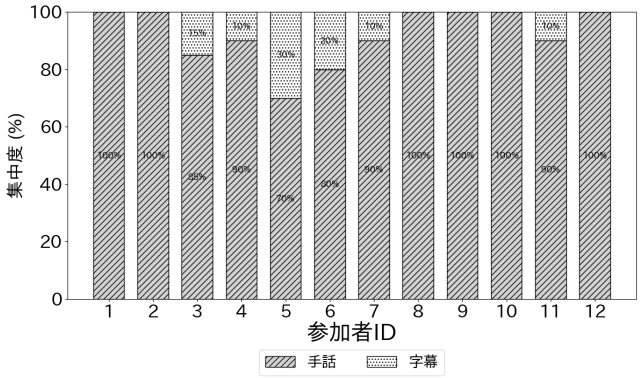

図 5　字幕提示手法 a における字幕と手話への集中度分布

参加者から報告された主な理由は以下の通りである．

- 手話への集中の必要性：多くの参加者が「手話の情報が最優先」「手話に集中してしまう」と報告している．
- 字幕による干渉：「字幕は邪魔だった」という意見もあり，一部の参加者にとっては字幕が手話の視聴を妨げる要因となっていた可能性がある．

これらの結果から，字幕提示方法 a を使用した場合，参加者の集中は圧倒的に手話に向けられており，字幕の優先順位が下がっていることが明らかになった．

4.4.3.2　AR 空間を活用した字幕提示方法 b,c での集中度

AR 空間を活用した字幕提示手法 b, c における字幕と手話への集中度の分布を図 6 に示す．

- 手話への集中強化：字幕が見えないことで，かえって手話への集中度が高まったという報告が複数挙げられた．

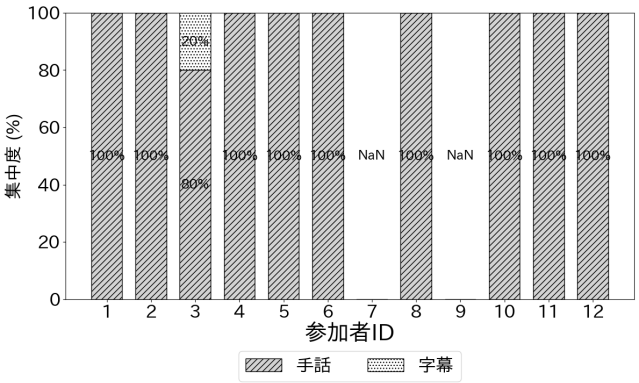

図 6　AR 空間を活用した字幕提示手法 b, c における字幕と手話への集中度分布

- 手話の視聴体験：AR 空間を活用した字幕提示手法 b, c を試した際の体験が AR グラス非装用時の手話視聴体験に近いものであったと報告している．

5.　考察

本研究では，AR グラスを用いた字幕提示システムの効果を検証するため，3 つのリサーチクエスチョン（RQ）を設定し，実験を行った．ここでは，各 RQ に対応する実験結果を基に考察を行う．

5.1　RQ1: 動画コンテンツの内容把握

RQ1「動画コンテンツの内容を把握することができるか」については，表 4 に示した試合内容理解度テストのスコアを用いて評価を行った．

結果として，字幕提示方法間で有意な差は認められなかった．この結果から，字幕提示方法による試合内容の理解度への影響が小さいと考えられる．さらに，表 7 に示されているように，字幕提示方法 c において，参加者が主観的に試合内容を把握できたと感じた平均が字幕提示方法 a よりも高い傾向がある．また，効果量も大きく，実質的な差異が存在することが示唆された．この結果は，ユーザー自身が字幕の提示位置をカスタマイズできることが，試合内容の理解度向上に寄与する可能性を示唆している．

したがって，AR グラスを用いた字幕提示システムにおいては，ユーザーが自由に字幕位置を調整できる機能を実装することが望ましいと考えられる．

5.2　RQ2: 対人コミュニケーション（手話）の内容把握

RQ2「対人コミュニケーション（手話）の内容を把握することができるか」については，表 6 に示した手話内容理解度テストのスコアを用いて評価を行った．

結果として，提示方法 a において，字幕内容誤答グループは字幕内容正答グループよりも有意に高い手話内容理解度を示した．この結果は，提示方法 a では字幕が手話の視認性を妨げている可能性を示唆している．

© 2024 Information Processing Society of Japan

一方，字幕提示方法 b，c では両グループ間に有意な差は見られなかった．このことから，AR 空間上での字幕提示方法（b，c）は，手話の視認性を維持しつつ，字幕からの情報取得を可能にする効果があると考えられる．ただし，表 5 に示されているように，手話動画間で理解度に有意差が認められたことから，手話動画の内容や難易度が結果に影響を与えた可能性がある．

今後は，手話動画の内容を統制した上で，字幕提示方法の効果を検証する必要がある．

5.3 RQ3: 各字幕提示方法の利点と課題

RQ3「各字幕提示方法の利点と課題はなにか」については，7 段階リッカート尺度でのフィードバックおよび実験全体終了後のフィードバックを用いて評価を行った．

字幕提示方法 a については， 4.3.2 節の自由記述形式のフィードバック解析から，字幕が視界の中心に固定されることへの不満が見られた．一方で，テレビと字幕の位置が同じであることによる見やすさも指摘された．字幕提示方法 b では， 4.3.2 節のフィードバック解析から，字幕と手話での情報提示が明確に区別されていたことが評価された．ただし，字幕位置のカスタマイズ性への要望は依然として高かった．字幕提示方法 c は， 4.3.2 節のフィードバック解析によると，字幕位置を自由に設定できることが高く評価された．字幕と映像の同時視聴が容易になったという意見もあった．ただし，微調整の必要性も指摘された．

図 4 に示されている実験全体のフィードバックでは，字幕提示方法 c が最も好ましいと評価された．これは，字幕位置の自由度が高いことが利用者のニーズに合致していることを示唆している．図 5 と図 6 は，字幕提示方法 a と AR 空間を活用した字幕提示方法 b, c とのそれぞれにおける手話と字幕の同時受容に関する結果を示している．これらの結果から，どちらの字幕提示方法であっても，手話が提示された際には字幕情報より手話情報を優先して情報を得ようとしていることが考えられる．

またフィードバック分析から，字幕提示位置と手話表現されている位置が重なってしまうと非常に邪魔に感じられることが明らかになった．このことから，手話と字幕の空間的配置に十分な注意を払う必要があるだろう．

6. まとめと今後の課題

本研究の結果から，字幕提示方法 a，b，c において，動画コンテンツでは表 4 から有意な差は見られなかった．しかしながら，図 4 より，AR 空間上でユーザーが字幕位置を自由に設定できる提示方法 c が，主観的評価において最も高い支持を得たことは注目に値する．このことは，Vinoba Vinayagamoorthy ら [5] が指摘するパーソナライゼーションの重要性を裏付けるものであり，字幕提示システムの設計においてもユーザーの個別のニーズや嗜好に対応することが不可欠であることを示唆している．

一方で，表 8 から字幕と手話の同時把握が提示方法によらず困難であるという課題も浮き彫りになった．この問題に対処するためには，視線の動きに合わせて字幕をゆっくりと追従させるなど，より直感的な操作を可能とする字幕提示方法が求められるだろう．また，字幕の内容や表示方法をユーザー個人が最適化できるようにすることで，手話との情報の重複を最小限に抑え，よりスムーズな情報取得が可能になると考える．

しかし，本研究にはいくつかの限界も存在する．実験環境や手話動画の内容が限定的であったことから，より多様な場面やコンテンツを対象とした検証が必要である．また，参加者の属性や個人差，特に手話や e-sports への理解度については十分に考慮できていない．今後は，ろう・難聴者の多様性を踏まえた上で，大規模かつ長期的な調査を行うことが求められる．

これらの課題はあるものの，本研究の結果は，AR グラスを用いた字幕提示方法がろう・難聴者の情報アクセシビリティ向上に大きく寄与する可能性を示している．今後は，本研究の知見を生かしつつ，ユーザーの多様なニーズに柔軟に対応できるインクルーシブなシステムの開発を進めていきたい．

参考文献

[1] XREAL（2023） ”Display Modes”， air2 製品ページ， URL: https://www.xreal.com/jp/air2 (参照日: 2023 年 12 月 9 日)

[2] 西岡知之，皆川洋喜，内藤一郎，清水豊: 単眼非透過 HMD による字幕提示の講義場面での聴覚障害者への情報保障の効果， ヒューマンインタフェース学会論文誌， Vol.9， No.2， pp.35-40， 2007.

[3] 加藤伸子，篠崎有斗: 透過型 HMD を用いた聴覚障害者への情報保障の試み-博物館での利用を目指して， 日本科学教育学会年会論文集 43， 2019， 123-124

[4] 阪田大輔; 森田健太郎; 佐藤健哉. スマートグラスにおける AR を用いた音声コミュニケーション支援手法の提案. マルチメディア， 分散協調とモバイルシンポジウム 2017 論文集， 2017， 2017: 1132-1137.

[5] Vinayagamoorthy, V. et al. (2019). Personalising the TV experience using augmented reality: An exploratory study on delivering synchronised sign language interpretation. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1-12.

[6] Berke, L. et al. (2019). Preferred Appearance of Captions Generated by Automatic Speech Recognition for Deaf and Hard-of-Hearing Viewers. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Paper LBW1713, 1–6.

[7] Crabb, M. et al. (2015). Online News Videos: The UX of Subtitle Position. In Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility, 215–222.

© 2024 Information Processing Society of Japan