

筑波大学 情報学群 情報メディア創成学類

卒業研究論文

表情に基づく動画再生速度の個人最適化による  
外国語学習支援

西田 直人

指導教員 志築 文太郎, 川口 一画

2022年2月

## 概要

外国語の学習には音読、精読、多読、精聴、多聴など、様々な学習方法があり、そのための教材としても本、動画、音楽など様々な媒体が用いられている。その中でも、海外ドラマなどの外国語動画をたくさん視聴する学習法（Extensive Viewing, 多視聴）には、母語話者の話す速度に慣れることができる、娯楽性が高く学習者のモチベーションを維持しやすい、劇中において実践的な言語の言い回しが用いられているなどの様々なメリットがある。しかし、多くの外国語学習者は、母語話者と同じように聴き取りを行えないために、多視聴を行う際、動画の再生速度についていけない場面に多く直面する。このため、母語話者向けの動画コンテンツを通常の再生速度のまま教材として扱うことは難しい。よって、本研究では動画再生速度の個人最適化システムの開発および有用性の検証を行った。結果から、本システムは全ての実験参加者の作業負担を高めず、英語が中級レベル（CEFR スコア：～B1 下位）の参加者において、内容に対する理解の向上が有意に見られた。すなわち、本システムを用いることにより、多聴学習や Extensive Viewing に代表される外国語の動画を視聴する学習方法において、個人のリスニング能力に最適化された速度にて動画を視聴することができることが分かった。本研究においては、学習者のシチュエーションコメディ視聴中の笑いどころにおける表情に基づき再生速度を最適化する。このシステムを用いて学習することにより、ビデオ教材の難易度を調整することができるため、多視聴学習を行う際に用いることができる学習教材の選択肢が広がる。

# 目次

<b>第1章</b>	<b>はじめに</b>	<b>1</b>
1.1	背景	1
1.2	目的とアプローチ	3
1.3	貢献	3
1.4	本論文の構成	3
<b>第2章</b>	<b>関連研究</b>	<b>5</b>
2.1	多聴および多視聴	5
2.2	既存の外国語学習支援システム	5
2.3	動画コンテンツの再生速度の変更	6
2.4	動画コンテンツに対する理解度の調査	6
2.5	語学力の差が非言語的行動に与える影響	6
2.6	本研究の位置付け	7
<b>第3章</b>	<b>予備実験</b>	<b>8</b>
3.1	目的	8
3.2	機器	8
3.3	実験参加者	8
3.4	手順	8
3.4.1	予備実験後のアンケート	11
3.5	収集データ	14
3.6	結果	14
3.6.1	予備実験後のアンケート	18
3.6.2	理解度の違いが外見的特徴に現れるか	18
3.6.3	内容理解に関するクイズの妥当性	19
3.6.4	録画映像を観察し分かった点	19
3.7	考察	21
<b>第4章</b>	<b>実装に関する予備調査</b>	<b>22</b>
4.1	プロトタイプ	22
4.2	デモで得たフィードバックに対する考察	23
<b>第5章</b>	<b>表情に基づく動画再生速度の個人最適化システム</b>	<b>25</b>

5.1	実装環境 . . . . .	25
5.2	学習者の理解の指標について . . . . .	25
<b>第6章</b>	<b>本実験</b>	<b>28</b>
6.1	目的 . . . . .	28
6.2	機器 . . . . .	28
6.3	実験参加者 . . . . .	28
6.4	実験設計 . . . . .	30
6.5	手順 . . . . .	32
6.6	収集データ . . . . .	33
6.7	結果 . . . . .	33
6.7.1	実験前アンケート . . . . .	34
6.7.2	実験中アンケート：語彙テスト . . . . .	34
6.7.3	システムを用いて動画を見た際の再生速度の変更記録 . . . . .	34
6.7.4	動画内容に関するクイズ . . . . .	37
6.7.5	実験後アンケート . . . . .	40
6.7.6	NASA-TLXによる作業負荷の指標 . . . . .	42
6.7.7	SUSによるユーザビリティの評価 . . . . .	43
6.7.8	システムに対する意見 . . . . .	43
6.8	考察 . . . . .	45
<b>第7章</b>	<b>議論</b>	<b>46</b>
<b>第8章</b>	<b>今後の展望</b>	<b>47</b>
<b>第9章</b>	<b>おわりに</b>	<b>48</b>
	謝辞	49
	参考文献	50
<b>付録A</b>	<b>予備実験にて用いた各種書類</b>	<b>57</b>
A.1	実験後アンケート . . . . .	57
<b>付録B</b>	<b>本実験にて用いた各種書類</b>	<b>61</b>
B.1	実験前アンケート . . . . .	62
B.2	実験中アンケート：語彙テスト1 . . . . .	63
B.3	実験中アンケート：語彙テスト2 . . . . .	68
B.4	実験中アンケート：内容把握に関するクイズ1 . . . . .	74
B.5	実験中アンケート：内容把握に関するクイズ2 . . . . .	75
B.6	実験中アンケート：SUS . . . . .	76

B.7 実験後アンケート . . . . .	78
------------------------	----

# 目次

1.1 システムの概略図. 動画の笑いどころにて学習者が笑っていないと, 動画の再生速度が遅くなる. . . . .	2
3.1 PC と参加者の位置関係および姿勢の様子のイメージ図. . . . .	9
3.2 予備実験中の画面の様子. 著作権のため, モザイク加工を行っている. . . . .	9
3.3 実験中提示されるクイズの様子. . . . .	10
3.4 動画の構成. . . . .	10
3.5 ラベリング中の様子. 実験参加者は1回目の視聴時に理解できなかった箇所にてキーボードを押下する. . . . .	10
3.6 アンケートの結果の一部. a) 英語はあなたの母語ですか. b) 外国語の動画を視聴している, または外国語で行われる会議や授業に参加している時, 理解が追いつかないと感じたことはありますか. c) 母国語の動画を視聴している, または会議や授業に参加している時, 理解が追いつかないと感じたことはありますか. d) 日常生活で英語はどれくらい使いますか (時間/週). . . . .	12
3.7 外国語のコンテンツを視聴している時に困ることはなんですかという選択式アンケートに対する回答. . . . .	13
3.8 分からない時にする表情は次のうちどれですかという選択式アンケートに対する回答. . . . .	13
3.9 教材の動画はどれくらい理解できましたかというアンケートに対する回答. (1: 全く分からない, 7: 全て理解できた) . . . . .	13
3.10 問題文が表示されている時間の長さはどうでしたかというアンケートに対する回答. (1: 極端に短すぎる, 7: 極端に長すぎる) . . . . .	14
3.11 キャプチャしたフレームの一例. 上が理解できていない時で, 下が理解できている時である. . . . .	15
3.12 クイズの点数と TOEIC の点数の関係. Pearson の相関係数=0.784 (p 値=0.116). Spearman の相関係数=0.821 (p 値=0.089) . . . . .	15
3.13 クイズの点数とクイズの提示時間のアンケートの関係. Pearson の相関係数=0.765 (p 値=0.131). Spearman の相関係数=0.703 (p 値=0.185) . . . . .	16
3.14 クイズの点数と動画に対する理解度の自己評価. Pearson の相関係数=0.576 (p 値=0.309). Spearman の相関係数=0.730 (p 値=0.161) . . . . .	16

3.15	本研究にて用いた Action Unit が起こす表情変化および担当する筋肉の種類. a) 眼輪筋および眼窩部は頬を引き上げる. b) 大頬骨筋は唇の端を引っ張る. c) 頬筋は頬の皺やえくぼを作る. d) 上眼瞼挙筋, 眼輪筋および眼窩部の脱力はまばたきを生じる. . . . .	17
3.16	笑顔に関する AU が動画中の笑いどころ (黄色の縦線で示されている) において示す値を散布図に表した. 理解している際の点は赤, および理解していない際の点は青でプロットしている. なお, 予備実験で用いた映像には笑いどころが約 6000 フレーム目から 7700 フレーム目までしか存在しなかったため, 5000 フレーム目から 8000 フレーム目までを描画範囲としている. . . . .	20
4.1	速度変更の実装のイメージ図. a) 笑いどころにおいて笑っていない場合, 速度が遅くなる. b) 笑いどころにおいて笑っていない場合, 一時停止が挿入される. . . . .	23
5.1	システムを使っている様子. 顔面部分の緑のフレームは無表情を示している.	26
5.2	表情検出についての説明. a) 緑のフレームは無表情を示す. b) 黄色の表情は笑顔を示す. . . . .	26
5.3	改良した速度変更の実装のイメージ図. 0.6 倍速から 1.0 倍速まで 0.1 倍刻みで速度が変わる. . . . .	26
5.4	Laughter-Detection ライブラリを用いて動画から笑いどころを抽出する工程図.	27
6.1	本実験にて, ドラマを見ている様子. . . . .	29
6.2	本実験にて, 問題提示画面を見ている様子. 実験参加者は手元の問題兼回答用紙の選択肢のうち正答と考えられるものをマークする. . . . .	29
6.3	統制の説明図. a) 動画の内容についての統制. b) 動画の構成. 9つのドラマのクリップおよびその内容に関するクイズから成る. c) グループのタスク統制. . . . .	30
6.4	事前アンケートの結果の一部. a) 英語はあなたの母語ですか. b) 日常生活で英語はどれくらい使いますか (時間/週) . . . . .	33
6.5	動画の経過時間に対し, 笑った点および笑っていない点の時刻をプロットした図. a) 動画 1 を見た者全員分のもの. b) 動画 1 を見た者のうち, TOEIC の点数が 700 未満の参加者のもの. c) 動画 2 を見た者全員分のもの. d) 動画 2 を見た者のうち, TOEIC の点数が 700 未満の参加者のもの. . . . .	35
6.6	グループ毎の再生速度の変化. a) A グループ. b) B グループ. c) C グループ. d) D グループ. . . . .	35
6.7	TOEIC のスコア毎の再生速度の変化. a) TOEIC の点数 800 以上で動画 1 を見た群. b) TOEIC の点数が 700 以上 795 以下で動画 1 を見た群. c) TOEIC の点数が 695 以下で動画 1 を見た群. d) TOEIC の点数 800 以上で動画 2 を見た群. e) TOEIC の点数が 700 以上 795 以下で動画 2 を見た群. f) TOEIC の点数が 695 以下で動画 2 を見た群. . . . .	36

6.8	AグループとDグループにおける動画内容に関するクイズの正答数と参加者のTOEICの点数の関係. AグループとDグループは動画1をシステムを用いて視聴し, 動画2をそのまま視聴した. . . . .	37
6.9	BグループとCグループにおける動画内容に関するクイズの正答数と参加者のTOEICの点数の関係. BグループとCグループは動画2をシステムを用いて視聴し, 動画1をそのまま視聴した. . . . .	38
6.10	動画1の内容に関するクイズの正答数と参加者のTOEICの点数の関係. AグループとDグループはシステムを用いて視聴し, BグループとCグループはそのまま視聴した. . . . .	38
6.11	動画2の内容に関するクイズの正答数と参加者のTOEICの点数の関係. BグループとCグループはシステムを用いて視聴し, AグループとDグループはそのまま視聴した. . . . .	39
6.12	事後アンケートの結果の一部. a) デバイスを用いた時に動画コンテンツを楽しむ事はできましたか. b) デバイスを用いなかった時に動画コンテンツを楽しむ事はできましたか. c) デバイスを用いた時の動画の速度調整は適切でしたか. d) 表情検出から自動で再生速度が変わるシステムを使用した際, 自分で速度を調整した方がいいと感じた場面がありましたか. . . . .	41
6.13	事後アンケートの結果の一部. a) よりよく理解できたのは動画1ですか動画2ですか (ただし, システム使用時/不使用時という可視化の仕方にした). b) 今回の動画コンテンツ (friends というドラマ) は以前に視聴したことがありますか. . . . .	41
6.14	NASA-TLX のスコアの平均値および 10% 標準誤差. . . . .	42

# 表目次

3.1	解析結果. それぞれの従属変数について, 学習者が理解していた条件下 (U: Understood) および学習者が理解していない条件下 (C: Confused) の独立変数の2群を比較している. . . . .	17
3.2	表情およびそれに対応する Action Unit. . . . .	18
6.1	TOEIC の点数とグループ分けの内訳. . . . .	31
6.2	動画前後の語彙テストや動画内容に関するテストの解答時間, ならびにシステムを用いた際の動画試聴時間. . . . .	32
6.3	語彙テストの点数. テストの点数の上限値が16点から18点と異なるため, 各セルに上限値を書いている. . . . .	34
6.4	動画内容に関するクイズの正答数と参加者の TOEIC の点数の相関関係の表. . . . .	37
6.5	動画内容に関するクイズの正答数の二群間検定の結果. . . . .	39
6.6	NASA-TLX のスコアの表. . . . .	42

# 第1章 はじめに

本論文は動画再生速度の個人最適化システムの開発および有用性の検証について述べたものである。本章では、本研究の背景、目的とアプローチ、貢献、および本論文の構成を述べる。

## 1.1 背景

外国語の学習方法の中に、多聴と呼ばれる学習方法がある。多聴とは、大量の外国語の言葉を聴くことである [1]。多聴による外国語学習には以下のような有用性がある。

- 聴いて理解する速度が速くなり、母語話者の話す速度に慣れることができる [2-6]。
- 耳で聴いて把握できる語彙が増える [2,5]。
- 2語以上の連結した発音による音声変化に慣れることができる [3,6]。
- 繋ぎ言葉、スラング、崩れた文法といった話し言葉のクセに慣れることができる [1,7]。

多聴は細かな文の分析をせず、全体的な文意を捉えることを重視する。

外国語の多聴学習には、音楽、ラジオ、または動画などが用いられる。その中でも特に、ドラマやアニメなどの外国語の動画コンテンツを視聴することにより多聴を行うことは、多視聴 (Extensive Viewing) と呼ばれている。多視聴による学習には、多聴の有用性に加え以下のような利点および特徴がある。

- 動画コンテンツは視覚情報があるため、音声のみのコンテンツと比べて内容理解を行うことが容易い [1,8]。
- 娯楽性が高い動画コンテンツであれば学習者のモチベーションを維持しやすい [9]。
- 母語話者向けの動画コンテンツであれば、劇中において実践的な言語の言い回しが用いられている [8,10,11]。
- 自分の決めた時間に行うことができる [8,12]。
- 字幕を用いれば、読む速度を向上させ、語彙知識を増やす学習にも効果的である [8,13-16]。
- 数ヶ月から数年といった長期間にわたり継続する必要がある学習方法である [1,8,17]。



図 1.1: システムの概略図. 動画の笑いどころにて学習者が笑っていないと, 動画の再生速度が遅くなる.

このように、多視聴は多くの有用性がある学習方法である。

しかし、多くの外国語学習者は、母語話者と同じように聴き取りを行うことが難しいために、多視聴を行う際、動画の通常再生速度において、話の流れについていけない場面に多く直面する [8, 18]。多聴や多視聴を行う際には、生徒の外国語能力に合わせた教材を選定する必要がある [1]。このため、母語話者向けの動画コンテンツを通常の再生速度のまま教材として扱うことは難しい。

その一方で、学習者が能動的に再生速度を変更して学習すると、学習者は再生速度を変更するための操作を煩わしく感じ、動画コンテンツを十分に楽しむことができず、学習に対するモチベーションが下がる可能性がある [19]。多聴や多視聴にとって、モチベーションの維持は学習効果に重要な因子である [1, 8, 17]。よって、モチベーション低下は避けたい課題である。

## 1.2 目的とアプローチ

動画コンテンツの再生速度という面から多聴の支援を行う目的に対し、学習者個人の理解度に応じて再生速度を自動で変動させるアプローチを取る。今回、予備実験の結果から、学習者個人の理解の有無を表情認識を用いて推定することとした。特に本研究では、笑いどころにて表情がこわばっていた場合、理解ができていないと判断した (図 1.1)。再生速度の調節方法については、学会のデモ発表におけるフィードバックを参考に、再生速度の減速を用いることにした。

## 1.3 貢献

本研究の貢献を以下に示す。

- 多視聴中の学習者のコンテンツに対する理解度を測るため、視線、表情、および行動に理解度の違いが現れていないか適切な指標を調査した。
- 動画コンテンツの難易度を学習者の表情に応じて再生速度を調節するシステムを提案し、実装した。
- 本システムの有用性を実験にて検証し、英語能力が中級 (CEFR スコアがおよそ B1 下位以下) である学習者が本システムを使用した際、内容に対する理解の向上が見られることを確かめた。

## 1.4 本論文の構成

第 1 章では、本研究の背景、本研究の目的とアプローチ、および本研究の貢献を示した。第 2 章では、本研究に関連する研究および手法を述べ、本研究の位置づけを行う。第 3 章では、

本手法の概要を述べる。第4章では、本手法の実装を行うために実施した、理解しているか否かによって表情、目線、および所作に違いが現れるかを調べた予備実験について述べる。第5章では、第4章の結果をうけ、本手法の実装方法に関して述べる。第6章では、実装した本手法に対し、評価実験を行うことにより本手法の性能を述べる。第7章では、本手法において現在わかっている問題点や改善点を述べる。第8章では、本手法の展望を述べる。第9章では、本研究の結論を述べる。

## 第2章 関連研究

以下では、本システムに関連する研究について述べる。

### 2.1 多聴および多視聴

多視聴は多聴の一種であり [1], 大量の外国語の動画コンテンツを視聴することである。多視聴独自の特徴としては、映像情報が理解の手助けになること [1,8], および、アニメ、またはシチュエーションコメディ（以降、シットコム）などの娯楽性が高い動画を視聴する場合、学習者のモチベーションが維持しやすいこと [9,11] がある。多視聴を行う際の注意点として、語彙に関する以下の事項がある。

- 学習対象言語の基本的語彙を 2000 語から 3000 語知っておく必要がある [20–23].
- 動画内容の理解に必要な語彙を事前に学習すると良い [2,24,25].

後者のものについては特に、ドラマの 1 エピソードを用いた実験において、動画内に出現する単語を 10 個ほど事前学習するだけで理解に大きな差が生じた研究結果もある [26].

また、多視聴においてはキャプション（学習対象言語の字幕）を併用すると良い [8,9,16,27,28]. キャプションを用いる利点としては、自分のリスニング語彙を超える動画もリーディング語彙を用いて見ることができ<sup>1</sup>、深い理解ができること [1,16,29], および、字幕情報がない動画コンテンツや母語の翻訳字幕がある動画コンテンツより、キャプションがある動画の方が語彙力の増強に役立つこと [13,24,28,30], 内容理解が進むことでモチベーションが高まること [9] が挙げられる。多視聴の題材としてはドラマ、特にシットコムと呼ばれるコメディドラマがよく用いられる。ドラマを用いる利点としては、映画などの動画コンテンツよりも短く動画が区切られていること [8], 母語話者が好んで用いる実践的な言い回しが登場すること [10,11] および、異なるエピソードであっても同じ背景知識が用いられることが多いことが挙げられる [31,32].

### 2.2 既存の外国語学習支援システム

外国語の動画を視聴することによる外国語の学習の際、動画の内容を理解できなくなる原因としては以下の 2 つが考えられる。

---

<sup>1</sup>一般にリスニング語彙はリーディング語彙の一部と言われる [1]

- 語彙や背景知識が分からない [33,34].
- 動画速度についていけない [35].

このうち、語彙が分からないという課題に関しては実装による支援方法が提示されてきた [19,33,34,36]. しかし、再生速度の側面から動画内容の理解を補助するシステムを実装した先行研究は少ない. よって本研究においては再生速度を調節する側面より、外国語学習支援を行う.

## 2.3 動画コンテンツの再生速度の変更

教材の再生速度について、教育学の分野にて研究が行われてきた. Blau [18] は、リスニング教材の再生速度を遅くする、または一時停止時間を設けることにより、動画内容に関するテストの点数が向上することを示した. したがって本研究においても、再生速度を遅くする、または一時停止時間を設ける機能を検討し、プロトタイプを作成した.

## 2.4 動画コンテンツに対する理解度の調査

動画コンテンツに対する理解度の調査について、Fujii および Rekimoto [19] は、外国語の動画コンテンツを視聴している際の学習者の視線を用い、彼らの動画コンテンツに対する理解度を推測した. この研究においては、学習者の視線が字幕の方へ向いているほど、学習者の外国語の能力が低い傾向があるということが示されている. さらに Fujii および Rekimoto は、学習者の理解度に合わせて字幕中の単語の対訳を行う補助を行い、劇中に出現する単語についての語彙テストの結果が、動画の視聴前後の比較にて向上したことを示した. 本研究においては、視線の他に、表情、まばたき、または体の動作といった特徴を通じて学習者の理解の度合いを測れるかを予備実験にて調べる. また、多視聴は自宅学習を想定している. 高価な機材を用いず、どこでもシステムを利用できるようにするため、高価な視線計測装置を用いず、ウェブカメラで得た映像に特徴が測れるかを調べる.

## 2.5 語学力の差が非言語的行動に与える影響

外国語の習熟度が低いことと、外国語使用時の緊張度 (Language Anxiety と呼ばれる) が高くなることは相関関係にある [37–40]. また、外国語使用時の緊張度が高いと、非言語的行動に差が出る [41]. この事実を確かめるために、本研究の予備実験において、実験参加者の外国語の習熟度と非言語的行動の関係を調べる. さらに、同一参加者の動画視聴中においても、理解ができている箇所と理解できていない箇所にて非言語的行動に違いが現れるか調べる.

## 2.6 本研究の位置付け

先行研究において、多視聴の際の動画再生速度の調節を行うための予備調査やシステム開発はされてこなかった。我々の研究では、学習者が動画コンテンツを理解できている箇所および理解できていない箇所の比較において、視線だけでなく、表情の変化、瞬きの回数、または貧乏ゆすりの有無など、身体的特徴の違いがウェブカメラを通じた情報に現れるか実験を通じて調査した。さらに、この結果をもとに表情に基づく再生速度調節機能を試作した。

## 第3章 予備実験

表情および視線といった学習者のコンテンツに対する理解度を測る適切な指標を選定するため、予備実験を実施した。本章において、その調査内容および結果に関して述べる。

### 3.1 目的

予備実験の目的は、外国語の動画を視聴している最中に、理解の有無により表情、視線、まばたき、および行動といった外見的特徴が現れるかを調べることである。

### 3.2 機器

予備実験には、ウェブカメラを内蔵したノート PC (Macbook Pro 13 inch, Big Sur バージョン 11.6, 2.8 GHz クアッドコア Intel Core i7, メモリ 16 GB) を用いた。PC を用いて、外国語の動画を参加者が視聴している最中、ウェブカメラの映像を録画した (図 3.1)。実験参加者の顔が鮮明に見えるように、実験に用いた部屋の照明とウェブカメラの位置関係について、逆光になるのを避けた。PC ディスプレイに対し参加者の顔の距離は PC 操作時の標準的な距離である 40 cm から 60 cm であった [42]。PC ディスプレイと参加者の顔が正対するように PC を配置した。

### 3.3 実験参加者

予備実験には、5 名 (平均 23.6 歳, 標準偏差 1.82, 全員男性) が参加した。5 名全員が英語を第二外国語として学んでいたため、英語を学習対象言語として選定した。実験参加者には実験参加前に、マスクを外す、および前髪を掻き分けるなど表情を見えやすくすることを指示した。また、普段の目の状態にて実験に参加するように指示した (裸眼 1 名, メガネ 2 名, コンタクト 2 名)。

### 3.4 手順

参加者はまず、ウェブカメラによって顔のフレーム画像を 30fps にて撮影されながら、5 本の母語話者向け英語動画コンテンツを英語音声および英語字幕の条件 (多視聴の条件) にて視聴し

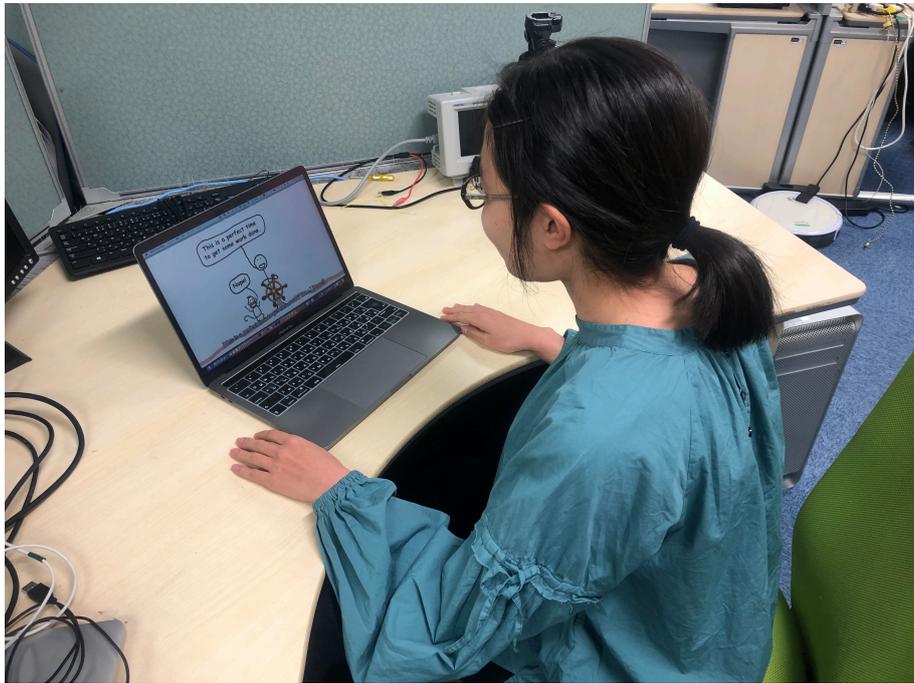


図 3.1: PC と参加者の位置関係および姿勢の様子のイメージ図.

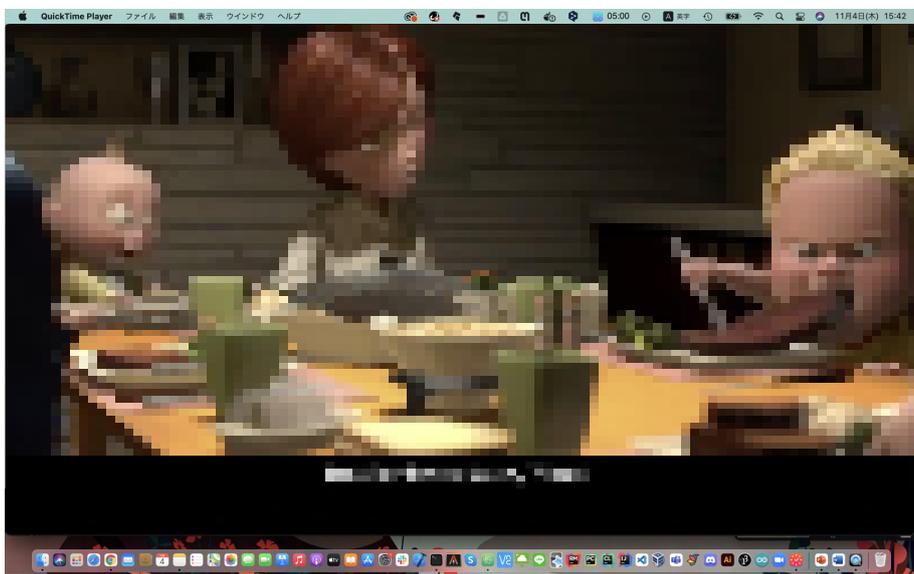


図 3.2: 予備実験中の画面の様子. 著作権のため, モザイク加工を行っている.

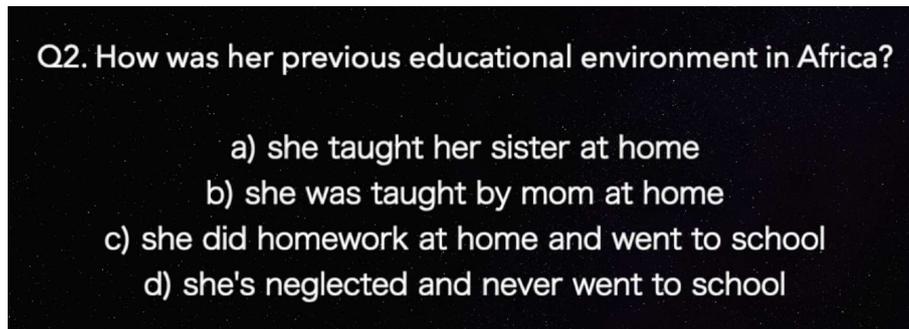


図 3.3: 実験中提示されるクイズの様子.



図 3.4: 動画の構成.



図 3.5: ラベリング中の様子. 実験参加者は1回目の視聴時に理解できなかった箇所にてキーボードを押下する.

た(図 3.2)。動画の長さは合計9分12秒であった。英語の動画コンテンツは、英語学習書 [43] に掲載されている TED Talks<sup>1</sup>, TED Education<sup>2</sup>, および洋画のクリップ<sup>3</sup>からおおよそ60秒から90秒ずつ切り取ったクリップを繋げて作成した。参加者の動画内容に対する理解度を大まかに知るために、動画コンテンツ中に、動画の内容に関する4択クイズをクリップとクリップの間に提示されるシーンを設けた(図 3.3, 図 3.4)。実用英語検定, TOEIC Listening&Reading, および TOEFL iBT のリスニング解答時間<sup>4</sup>を参考にし、4択クイズは1問あたり15秒提示された。動画内容に関する4択クイズを出題することにより、動画内容の把握具合を調べることができるとともに、実験参加者の意識を動画に対し集中させることができる。実験参加者はもう一度動画コンテンツを視聴しつつ、フレーム画像に対し、理解ができていた箇所および理解できていなかった箇所のラベリングを行った(図 3.5)。理解ができていた基準としては [8] の基準に従い、瞬時に大まかに文意が捉えられることを理解の基準とした。また参加者には実験後に、参加者自身のことについて、および実験中の感想や行動についてのアンケートに回答して貰った。実験参加者は実験中図 3.1 のような姿勢を続けた。実験中に顔の位置を大きく動かしたものはなかった。実験の所要時間は1人当たり約30分であった。

### 3.4.1 予備実験後のアンケート

予備実験後のアンケートの項目について、以下に示す。

- 英語はあなたの母語ですか(図 3.6 a)
- 外国語の動画を視聴している、または外国語で行われる会議や授業に参加している時、理解が追いつかないと感じたことはありますか(例:学部の英語の授業、洋画の視聴、国時会議のプレゼン動画の視聴、国時会議への参加時)(図 3.6 b)。
- 母国語の動画を視聴している、または会議や授業に参加している時、理解が追いつかないと感じたことはありますか(例:難しい数学の授業)(図 3.6 c)。
- 日常生活で英語はどれくらい使いますか(時間/週)(図 3.6 d)。
- 外国語のコンテンツを視聴している時に困ることはなんですか(自由記述式, 図 3.7)。
- 分からない時にする表情は次のうちどれですか(図 3.8)。
- 教材の動画はどれくらい理解できましたか(図 3.9)。
- 問題文が表示されている時間の長さはどうでしたか(図 3.10)。

<sup>1</sup><https://www.youtube.com/watch?v=arj7oStGLkU>, <https://www.youtube.com/watch?v=eIho2S0ZahI>

<sup>2</sup><https://www.youtube.com/watch?v=eM2VWspRpfk>

<sup>3</sup><https://www.youtube.com/watch?v=Rnwwo9Zol6w>, <https://www.youtube.com/watch?v=LZ0v9A5iEG4>

<sup>4</sup>実用英語検定のリスニング解答時間は10秒(<https://www.eiken.or.jp/s-cbt/test/2020.html>)。TOEIC のリスニング解答時間はおよそ5秒から20秒(<https://eigo.plus/toeic/toeic-9>)。TOEFL iBT のリスニング解答時間は15秒から30秒 [44]。

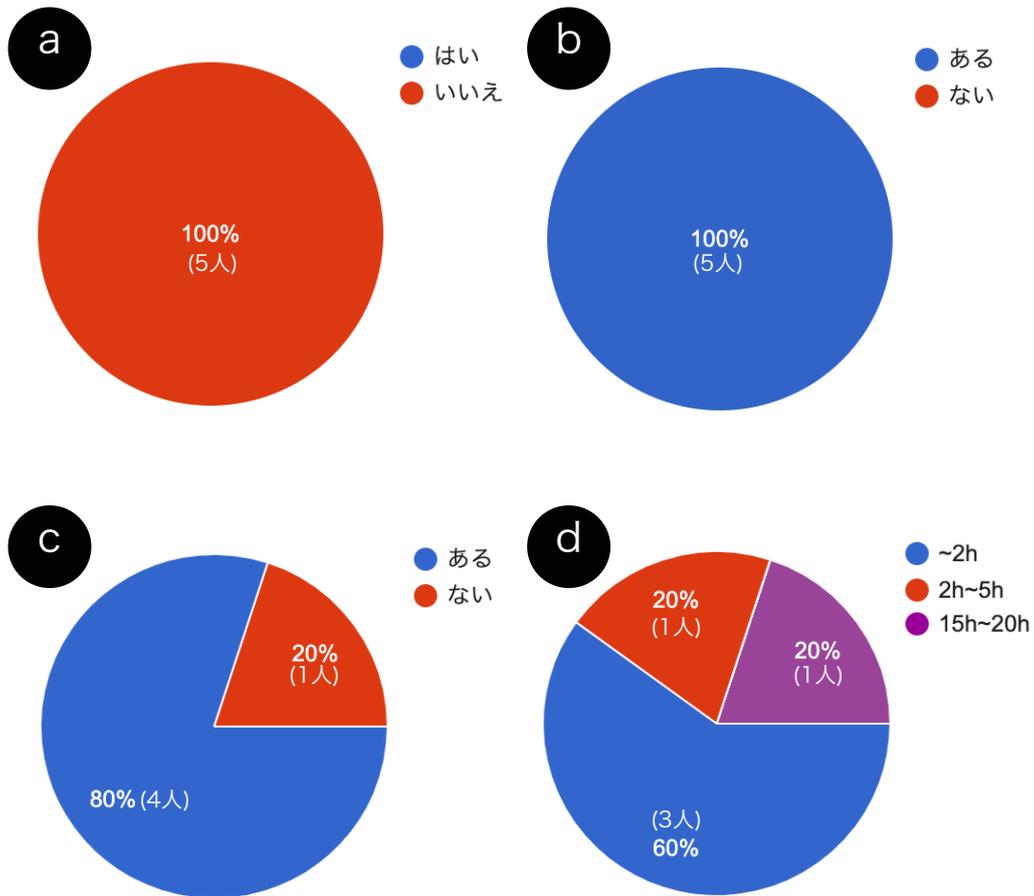


図 3.6: アンケートの結果の一部. a) 英語はあなたの母語ですか. b) 外国語の動画を視聴している, または外国語で行われる会議や授業に参加している時, 理解が追いつかないと感じたことはありますか. c) 母国語の動画を視聴している, または会議や授業に参加している時, 理解が追いつかないと感じたことはありますか. d) 日常生活で英語はどれくらい使いますか (時間/週).

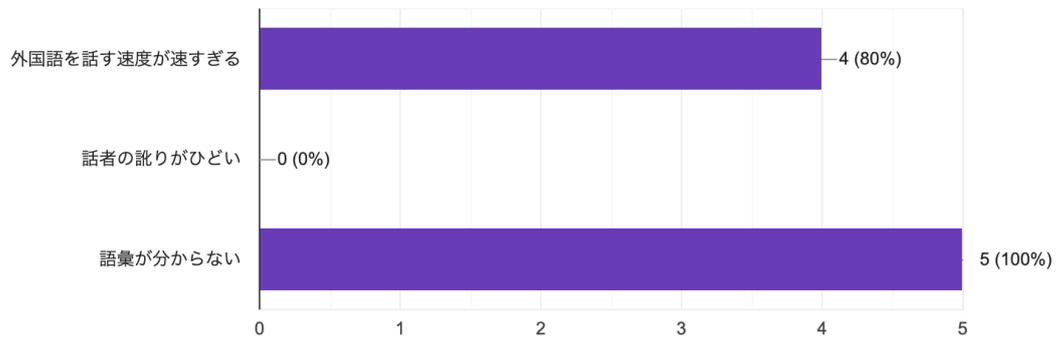


図 3.7: 外国語のコンテンツを視聴している時に困ることはなんですかという選択式アンケートに対する回答.

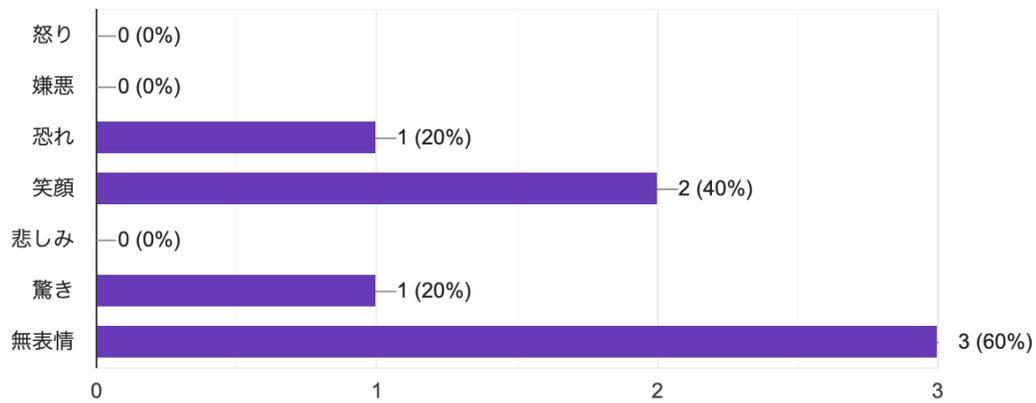


図 3.8: 分からない時にする表情は次のうちどれですかという選択式アンケートに対する回答.

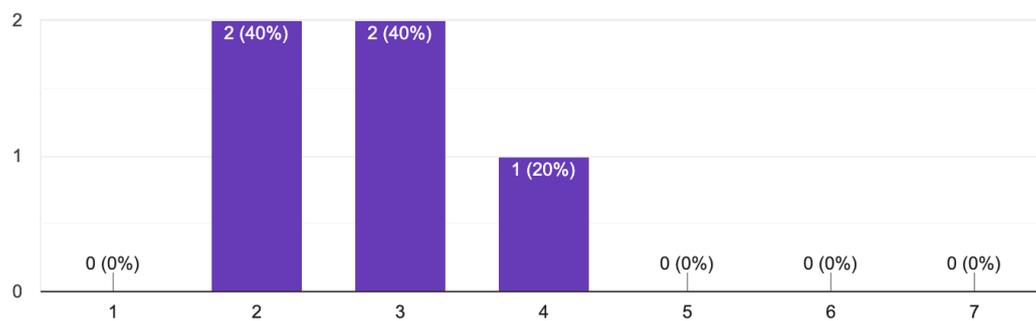


図 3.9: 教材の動画はどれくらい理解できましたかというアンケートに対する回答. (1: 全く分からない, 7: 全て理解できた)

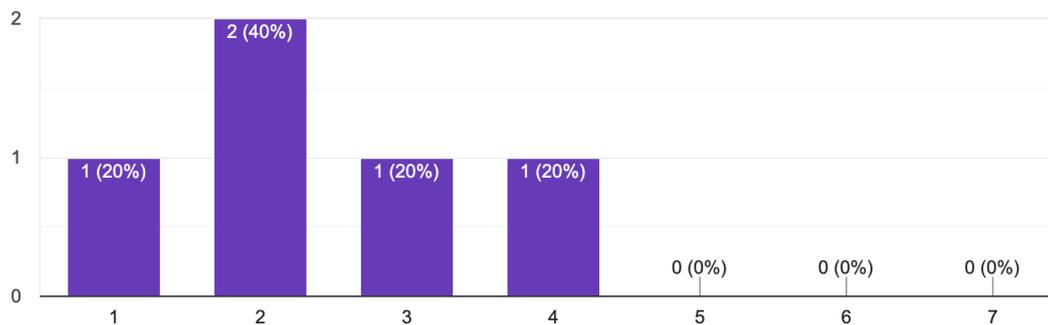


図 3.10: 問題文が表示されている時間の長さはどうでしたかというアンケートに対する回答。(1: 極端に短すぎる, 7: 極端に長すぎる)

- ビデオの中において理解が難しい場面はありましたか (印象に残る場面だけでも)。
- 分からない (自分の理解が追いつかない) 時に取る癖はありますか。ある場合は軽く記述してください (貧乏ゆすり, 視線が逸れる, 表情が変化するなど)。
- 外国語能力の指標を教えてください (例: TOEFL, TOIEC, 英検などのスコア, CEFR スコア, もしくは体感でどの動画が何%ほど理解できるか)。

### 3.5 収集データ

実験にて収集し, 解析に用いたデータは以下の通りである。

- 実験中のウェブカメラの録画映像
- カメラの録画映像をフレームに分割し, 参加者の理解 (理解している / 理解していない) によってラベル付けしたもの (約 13230 枚 / 人, 24fps)
- 実験中の動画の内容に関する 4 択クイズに対する解答
- 実験後アンケート (付録 A)

ウェブカメラの録画映像は, サッカーカードおよび停留点といった時系列に依存する視線の特徴をウェブカメラにて取得するために用いた。理解具合によりラベル付けされたフレーム画像は, 参加者の視聴中の理解度合いが外見的特徴に変化を及ぼすかを調べるために用いた。

### 3.6 結果

以下では, 予備実験にて取得したそれぞれのデータの解析を行った結果を示す。



図 3.11: キャプチャしたフレームの一例. 上が理解できていない時で, 下が理解できている時である.

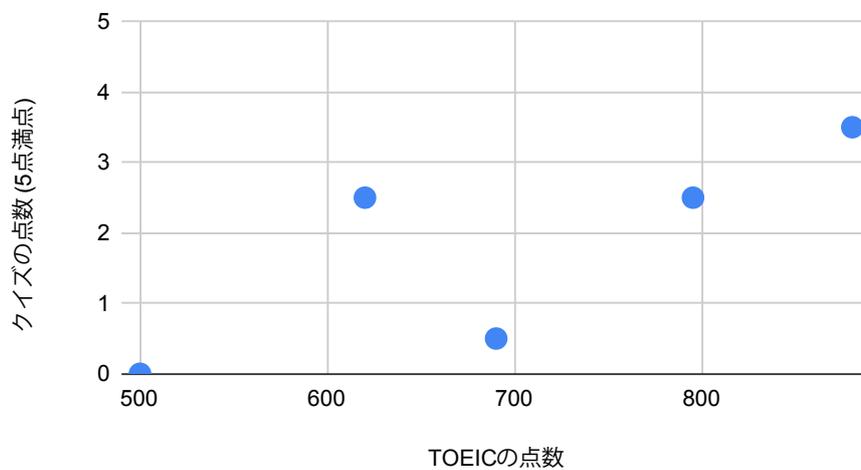


図 3.12: クイズの点数と TOEIC の点数の関係. Pearson の相関係数=0.784 (p 値=0.116). Spearman の相関係数=0.821 (p 値=0.089)

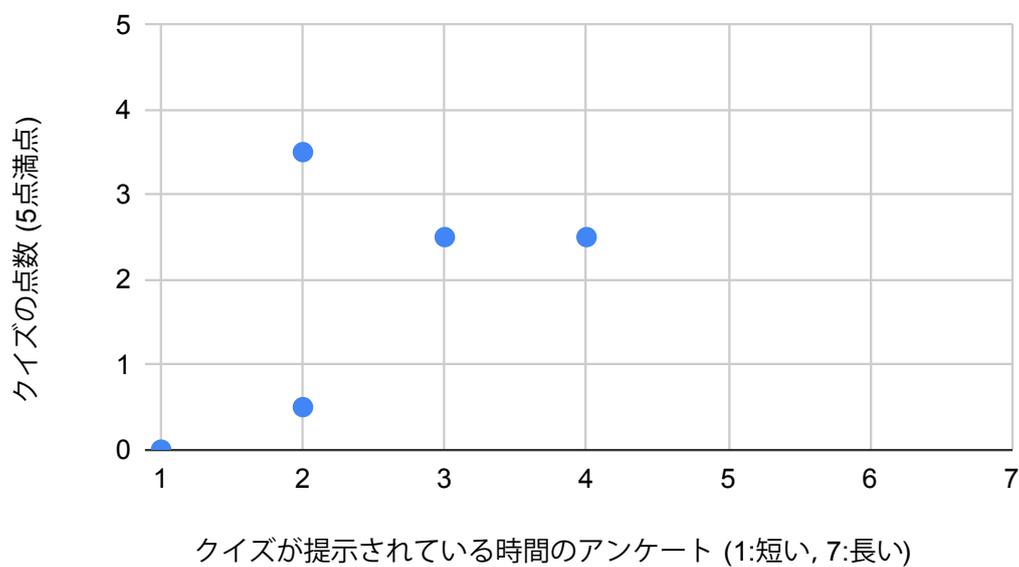


図 3.13: クイズの点数とクイズの提示時間のアンケートの関係. Pearson の相関係数=0.765 (p 値=0.131). Spearman の相関係数=0.703 (p 値=0.185)

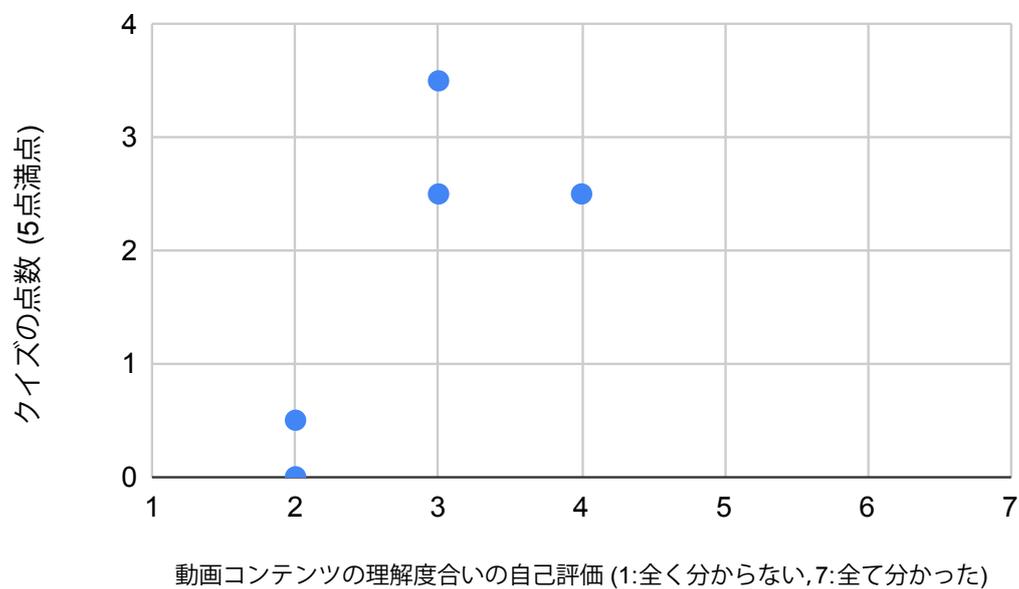


図 3.14: クイズの点数と動画に対する理解度の自己評価. Pearson の相関係数=0.576 (p 値=0.309). Spearman の相関係数=0.730 (p 値=0.161)



図 3.15: 本研究にて用いた Action Unit が起こす表情変化および担当する筋肉の種類. a) 眼輪筋および眼窩部は頬を引き上げる. b) 大頬骨筋は唇の端を引っ張る. c) 頬筋は頬の皺やえくぼを作る. d) 上眼瞼挙筋, 眼輪筋および眼窩部の脱力はまばたきを生じる.

表 3.1: 解析結果. それぞれの従属変数について, 学習者が理解していた条件下 (U: Understood) および学習者が理解していない条件下 (C: Confused) の独立変数の 2 群を比較している.

検定値 (理解している: U, 理解していない: C)	中央値		平均値 (標準偏差)		有意確率	効果量	効果量 95%信頼区間	
	U	C	U	C			下限	上限
Gx	-0.001	-0.006	0.008 (0.081)	0.000 (0.064)	p < .001	0.097 (無視可能)	0.076	0.118
Gy	0.293	0.307	0.291 (0.096)	0.307 (0.083)	p < .001	-0.169 (無視可能)	-0.190	-0.148
Gx × Gy	-0.000	-0.002	0.002 (0.024)	0.002 (0.020)	p < .001	0.011 (無視可能)	-0.010	0.032
AU6	0.100	0.230	0.389 (0.554)	0.387 (0.563)	p < .05	0.003 (無視可能)	-0.018	0.024
AU12	0.040	0.000	0.331 (0.466)	0.295 (0.494)	p < .001	0.076 (無視可能)	0.055	0.097
AU14	0.000	0.000	0.254 (0.394)	0.170 (0.290)	p < .001	0.222 (小)	0.201	0.243
AU6 × AU12	0.000	0.000	0.319 (0.873)	0.310 (1.067)	p < .001	0.009 (無視可能)	-0.012	0.030
AU6 × AU14	0.000	0.000	0.221 (0.499)	0.155 (0.436)	p < .001	0.136 (無視可能)	0.115	0.157
AU12 × AU14	0.000	0.000	0.201 (0.480)	0.138 (0.383)	p < .001	0.136 (無視可能)	0.115	0.157
AU6 × AU12 × AU14	0.000	0.000	0.234 (0.896)	0.183 (0.854)	p < .001	0.057 (無視可能)	0.036	0.078
AU45	0.000	0.000	0.188 (0.372)	0.199 (0.382)	p > .03	-0.031 (無視可能)	-0.052	-0.010

表 3.2: 表情およびそれに対応する Action Unit.

表情	対応する Action Unit	表情に対しユニークな Action Unit もしくはその組み合わせ
驚き	AU 1, 2, 5, 15, 16, 20, 26	AU 16
恐れ	AU 1, 2, 4, 5, 15, 20, 26	AU (4, 5) <sup>†</sup>
不快	AU 2, 4, 9, 15, 17	AU 17
怒り	AU 2, 4, 7, 9, 10, 20, 26	AU 10
幸福	AU 1, 6, 12, 14	AU 6, 12, 14
悲しみ	AU 1, 4, 15, 23	AU 23

<sup>†</sup> AU4 および AU5 の組み合わせという意味である。

### 3.6.1 予備実験後のアンケート

実験参加者全員が英語を第二外国語として学んでおり、母語話者はいなかった。また、日常生活にて英語を用いる時間について、週に2時間以下という回答が大多数を占めていた。外国語の動画を視聴したり、外国語で行われる会議や授業に参加している時に理解が追いつかないと感じたことがあると全員が解答した。さらに、母国語を用いた動画、会議、および授業についても以上のような傾向があった。外国語のコンテンツを視聴している時に困ることとしては、語彙が分からないことを全員が問題としてあげており、外国語を話す速度が速すぎることもほとんどの人が問題として挙げていた。リッカート尺度を用いて計測した教材および内容理解の問題の難易度については、どちらも難しい傾向にあった。分からない時に取る癖や表情については、特に一貫したものは申告されなかった。

### 3.6.2 理解度の違いが外見的特徴に現れるか

実験後、参加者の顔のフレーム画像(図 3.11)を解析した。フレーム画像は全部で 62734 枚あり、動画の内容を理解できたトラベリングされたものが 52385 枚、理解していないトラベリングされたものが 10349 枚であった。顔の特徴量を取得するために、Openface ライブラリ [45] の視線追跡ユニット [46] および Action Unit (AU) 検出ユニット [47] を使用した。AU とは、人間の顔の筋肉部位に対応した 66 個の特徴点により表情をモデル化したものである [48,49]。視線追跡ユニットを用いて取得したデータは、 $G_x$  (視線のベクトルの x 成分) および  $G_y$  (視線のベクトルの y 成分) である。Action Unit 検出ユニットを用いて取得したデータは、AU1, AU2, AU4, AU5, AU6, AU7, AU9, AU10, AU12, AU14, AU15, AU17, AU20, AU23, AU25, AU26, および AU45 である。これらは Openface により取得できる全ての Action Unit である。それぞれの Action Unit がどの表情の指標になるかを表 3.2 に示す。解析には、正規性の確認のために Kolmogorov-Smirnov 検定を用いた。その結果、検定を行う値について、全て正規性が確認できなかった。よって、ノンパラメトリック検定のうち 2 群間の比較に用いられる Wilcoxon の順位和検定を用い、動画内容の理解ができてい

理解ができていない際に従属変数に差があるかを調べた。その結果、AU45以外において2群間に差が見られた。Hedges' gを用いて効果量を測定した結果、ほとんどの従属変数において無視できるほど小さかった。しかし、AU14において0.2220448と小さな効果量が検出された。AU14は笑顔に関連するAction Unitである[50,51]。よって笑顔に関連するAU (AU6, AU12, AU14)の効果量について詳しく調べた(表3.1)。それぞれのAUが示す筋肉部位の図を図3.15に示す[52]。なお、今回効果量の解釈には[53]を用い、「大(大きい)」「中」「小(小さい)」「無視可能」の4段階にて評価した。

### 3.6.3 内容理解に関するクイズの妥当性

内容理解に関するクイズの正答率については、0点正解が1人、0.5点正解が1人、2.5点正解が2人、3.5点正解が1人であった(0.5点正解は複数選択の問題があったため)。内容理解に関するクイズと実験参加者のTOEICの点数の関係は図3.12に示される。内容理解に関するクイズと実験参加者のTOEICの点数には正の相関(相関係数0.784)があるため、内容理解に関するクイズの点数は参加者の英語力を反映していると考えられる。ゆえに、内容理解に関するクイズの成績から、どの実験参加者も動画コンテンツの内容を部分的にしか理解していなかったと考えられる。内容理解に関するクイズの成績、クイズの提示時間、動画に関する理解度の自己評価についての関係は図3.13および図3.14に示される。内容理解に関するクイズの成績と動画に関する理解度の自己評価にはある程度の正の相関が見られた(相関係数0.765)が、内容理解に関するクイズの成績とクイズの提示時間にはそれほど相関関係は見られなかった(相関係数0.576)。

### 3.6.4 録画映像を観察し分かった点

また、アンケート結果を踏まえ、録画映像を観察し分かったこととして次の2点がある。

- 動画コンテンツ内の笑いどころに際し、参加者は内容を理解できる箇所において笑い、内容を理解できていない箇所において無表情を保った
- 参加者は動画コンテンツ内の笑いどころではない箇所において、無表情を保つ

予備実験後の自由記述アンケートにおいて、内容を理解できない際に笑いどころにおいても無表情を保つ理由としては、集中しようとして顔がこわばる、理解ができないから話の笑いどころが分からない、および、分からない箇所があまりに多いと表情を変えている余裕がないという意見があった。よって、笑顔に関するAUが理解度の有無によって示す値を笑いどころが出現する付近にて可視化した(図3.16)。この図から、理解ができていないと、理解している時よりもAUの値は小さくなる傾向にあることがわかった。以上の結果から、外国語の動画コンテンツへの理解が十分にできていない場合、参加者は動画コンテンツを楽しめず、動画内の笑いどころにおいても無表情を続ける傾向にあると分かった。

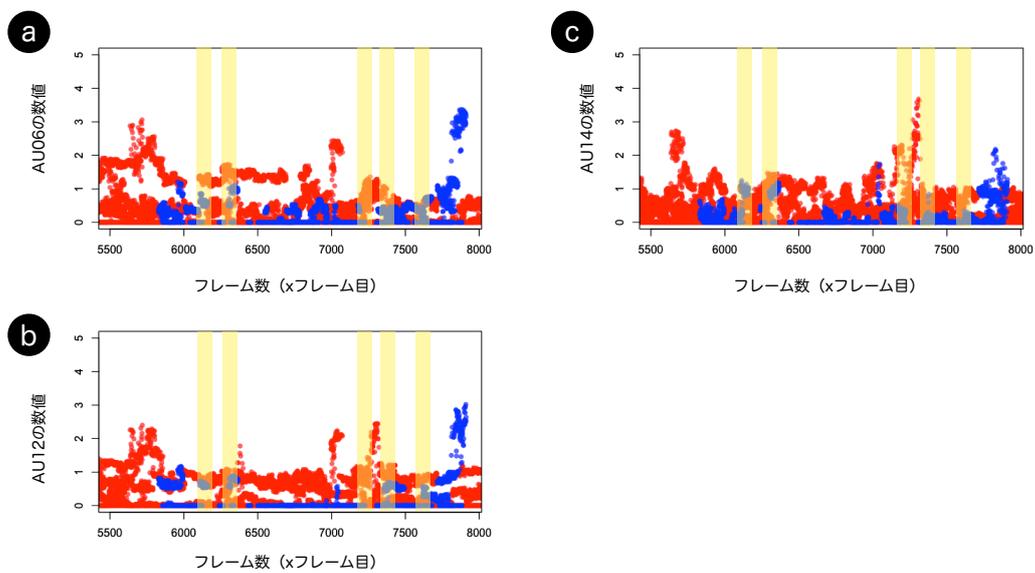


図 3.16: 笑顔に関する AU が動画中の笑いどころ（黄色の縦線で示されている）において示す値を散布図に表した．理解している際の点は赤，および理解していない際の点は青でプロットしている．なお，予備実験で用いた映像には笑いどころが約 6000 フレーム目から 7700 フレーム目までしか存在しなかったため，5000 フレーム目から 8000 フレーム目までを描画範囲としている．

### 3.7 考察

外国語の動画を視聴している最中に、理解の有無により外見に特徴が出るかを調べた結果、笑顔に差が現れることが分かった。また、録画映像を確認し、自由記述アンケートへの回答を踏まえた結果、話の理解ができていないと、話の笑いどころにおいても笑うことができず無表情になってしまうということが分かった。よって、必ずしも笑いどころで笑わないので話の理解ができていないとは言い切れないものの、笑いどころにおいて笑わず、無表情であるということは、話の理解ができていないことへのある程度の指標（必要条件）になると考えられる。

したがって本研究においては、動画中の笑いどころにて笑っていると理解ができており、動画中の笑いどころにて笑っていないと理解ができていないものとして考える。

## 第4章 実装に関する予備調査

予備実験の結果から、プロトタイプを実装した。国内会議のワークショップにてデモ発表を行った際、プロトタイプについて意見が得られたので本章に記す。

### 4.1 プロトタイプ

予備実験の結果より、学習者の理解度を測る指標として学習者が笑いどころにて笑っているか否かを用いることにした。

速度調節の方法については、[18]を元に、学習者が動画コンテンツの笑いどころにて笑っていない場合に、再生速度を調節する2種類の機能を筆者は実装した(図4.1)。

1つ目は、再生速度を遅く調節する機能である(図4.1a)。再生速度は先行研究[54]にならい、1倍速、0.8倍速、0.6倍速、および0.4倍速の4段階である。笑いどころにおいて笑っていないと1段階下がり、笑いどころにおいて笑っているとそのままの速度を維持する。

2つ目は、動画の再生途中に一時停止時間を設けることにより、学習者が理解しやすいように調節する機能である(図4.1b)。ただし、構文解析および音声解析の実装上の都合から、[18]と異なり、一時停止の挿入の実装については、文節の区切りにおいて一時停止時間を設けるものではなく、3秒ごとに0.5秒の一時停止時間を設けることにした。一回の一時停止の時間である0.5秒は、0.5秒程度の一時停止時間を設けることにより英文の復唱の再現性(シャドローイングの精度)が向上したという先行研究[55]に基づいている。一時停止の時間間隔である3秒は、人間の短期記憶がおおよそ3秒間持続するという研究結果[56]に基づいている。一時停止時間の挿入については、笑いどころから次の笑いどころまで持続し、次の笑いどころにおいて笑うと通常再生に戻り、笑っていないと一時停止が挿入されたまま再生される仕組みになっている。これらの実装から、調整後の動画コンテンツの笑いどころにて学習者が笑った場合、学習者が理解可能な速度に動画が自動調整されたと解釈できる。

多視聴に用いられる動画はシットコムが良いとされている[8-10]ため、再生する動画にはYouTubeのfriends(シットコムの代表的作品)の動画クリップを用いた<sup>1</sup>。YouTubeの動画を用いた理由として、動画の再生速度を変更した際にピッチ調整が自動で行われるため、再生速度の変更から来る違和感および不快感が生じにくいと考えたためである。

笑いどころの定義としては、シットコムの笑いどころにてしばしば流れる録音笑いが聞こえる箇所を笑いどころと定義した。実装としては、予め動画中にて笑い声が聞こえる箇所の

---

<sup>1</sup>[https://www.youtube.com/watch?v=nvzkHGndt\\_fk](https://www.youtube.com/watch?v=nvzkHGndt_fk)

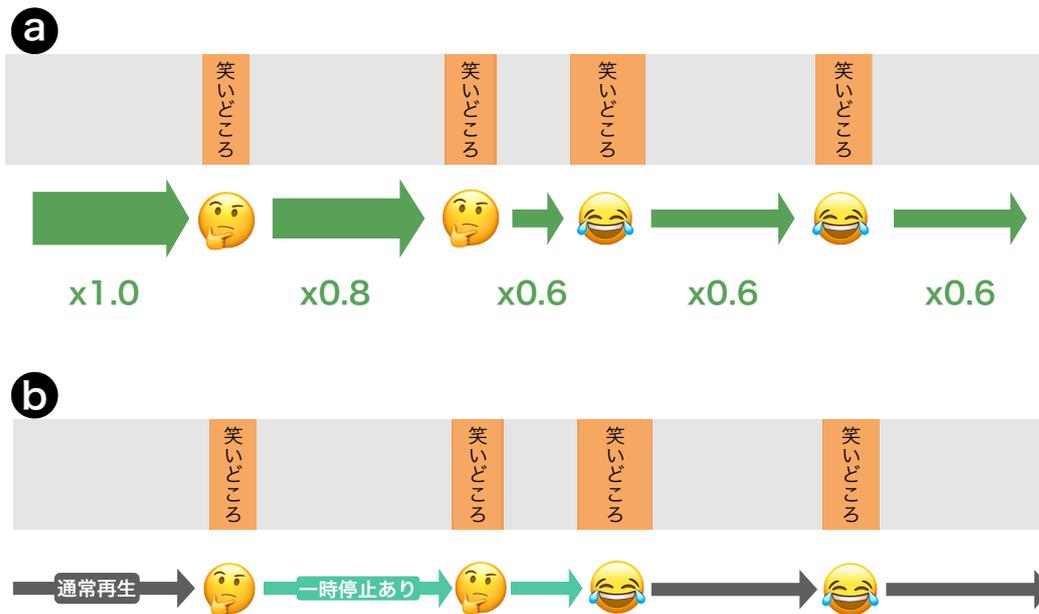


図 4.1: 速度変更の実装のイメージ図. a) 笑いどころにおいて笑っていない場合、速度が遅くなる. b) 笑いどころにおいて笑っていない場合、一時停止が挿入される.

タイムスタンプを Laughter-Detection ライブラリ [57] にて抽出し、筆者の手により手直しを施した.

学習者の笑顔を検出するための表情検出には、PAZ ライブラリ [58] の MiniXception モデル [59] を用いた.

## 4.2 デモで得たフィードバックに対する考察

第 29 回インタラクティブシステムとソフトウェアに関するワークショップにおいてデモ発表 [60] を行い、来場者およそ 20 人ほどに使用してもらった結果、以下のようなフィードバックを得られた.

1. 再生速度を遅くする実装は、なかなか気づかないほど自然に速度が変化する.
2. 再生速度を遅くする実装において、0.4 倍速になると音割れが気になる.
3. 再生速度を遅くする実装において、時々自分が聞き取れる速度よりも遅くなってしまいイライラする.
4. 一時停止時間を設ける実装は、回線速度が遅いように感じてしまい、あまり快適ではない.
5. コメディ動画という表情認識を活かせそうなチョイスが良い.

2. について、音割れが気になるということは、快適に視聴ができていないということである。これは、視聴者の認知負荷は極力下げるという多視聴の原理に反している。よって、音割れが起こらない速度変更の範囲に再生速度を収める必要がある。

次に、3. について、3. が起きる原因としては、学習者が理解できたが笑っていないというケースにおいて、システムの誤判断（偽陽性）が発生していることが考えられる。システムの誤判断によって余計に減速が行われるため、必要以上に遅くなってしまい、

最後に、4. について、4. が発生する原因としては、文節の区切りにて一時停止時間が設けられていないことが考えられる。プロトタイプの実装方法では逆に認知負荷が高まることが分かった。

## 第5章 表情に基づく動画再生速度の個人最適化システム

ワークショップにおいて受けたフィードバックをもとに、ユーザの表情を検出し動画の再生速度を調整するシステムを実装した。

### 5.1 実装環境

実装言語に Python3.7.4 を用いた。参加者の表情を検出するためにウェブカメラにて取得したフレーム画像および Perception for Autonomous Systems ライブラリ<sup>1</sup> [58] (図 5.1, 図 5.2), ブラウザ操作のために Selenium [61] を用いた。

また、再生する動画コンテンツについては、先行研究 [8,9] において調査されたシットコムの教材としての有用性から、デモで用いた friends を引き続き動画コンテンツの題材として用いることにした。また、YouTube 上にある動画を操作し、再生速度を調節することにした。理由としては、YouTube の再生速度変更機能にはタイムストレッチアルゴリズムが用いられており、速度変更を行った際にピッチ調整が行われ、快適に動画を視聴できるためである。

また、予め動画中にて笑い声が聞こえる箇所を Laughter-Detection [57] にて抽出し、その箇所を笑いどころと定義した (図 5.4)。

### 5.2 学習者の理解の指標について

第4章において「学習者が理解できたが笑っていない」場合が実際に生じたことを考える。「理解していない」という必要条件をトリガーとしているため、このシステムは偽陽性の問題を常にはらむ。よって、偽陽性が出た場合にも感づかれないような実装が好ましい。

第4章において「再生速度を遅くする実装は、なかなか気づかないほど自然に速度が変化する」「一時停止時間を設ける実装は、回線速度が遅いように感じてしまい、あまり快適ではない」という意見をもらった。したがって今回、再生速度の調節手法を再生速度を遅くする実装のみに限定し、実験を行うことにする。さらに、速度調節の段階について、1倍速、0.8倍速、0.6倍速、および0.4倍速の4段階から、1倍速、0.9倍速、0.8倍速、0.7倍速、0.6倍速の5段階に刻み幅を細かくした。これは、自然な速度変化をより実現するためである。下限を0.6倍速としたのは、前章における「再生速度を遅くする実装において、0.4倍速になる

<sup>1</sup>Openface ライブラリ [45] はリアルタイム処理を行うことができないため。

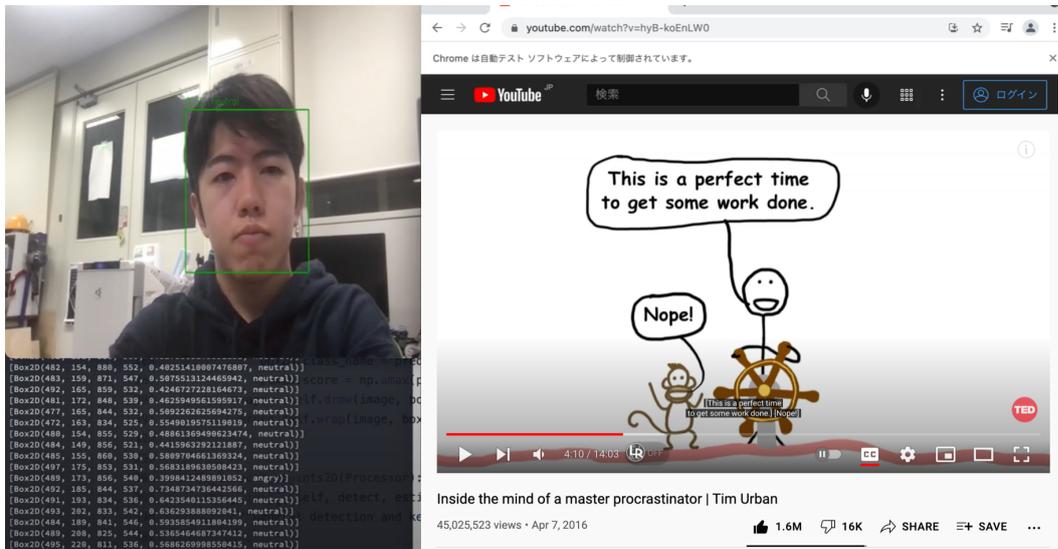


図 5.1: システムを使っている様子. 顔面部分の緑のフレームは無表情を示している.

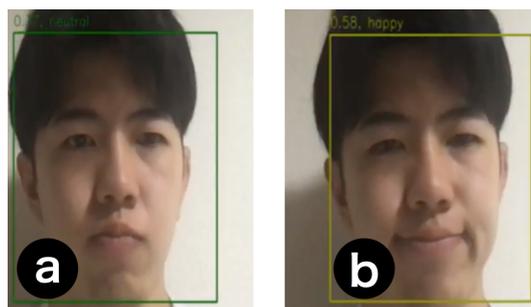


図 5.2: 表情検出についての説明. a) 緑のフレームは無表情を示す. b) 黄色の表情は笑顔を示す.

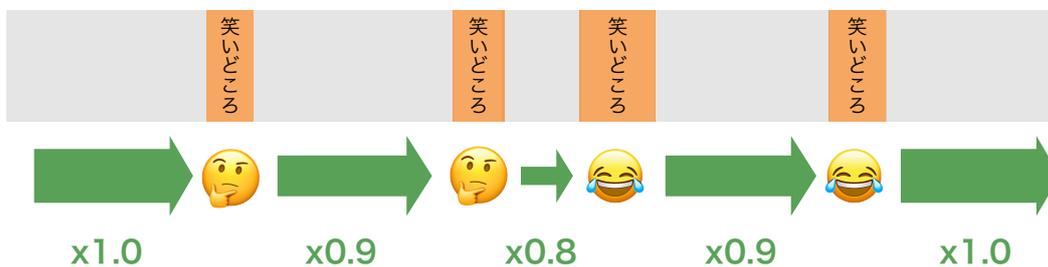


図 5.3: 改良した速度変更の実装のイメージ図. 0.6 倍速から 1.0 倍速まで 0.1 倍刻みで速度が変わる.

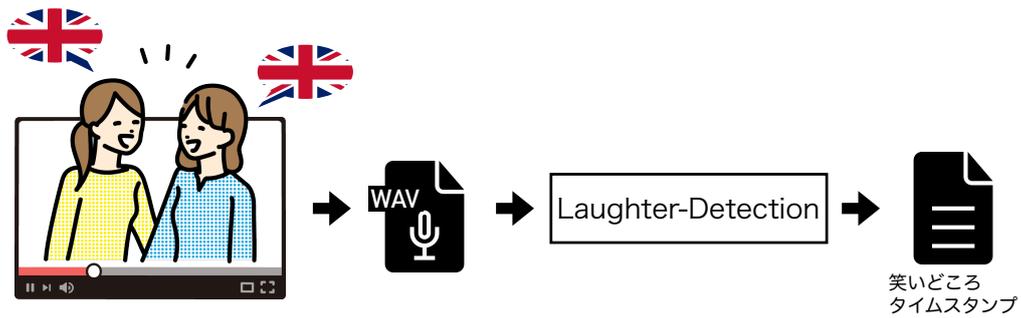


図 5.4: Laughter-Detection ライブラリを用いて動画から笑いどころを抽出する工程図.

と音割れが気になる」という意見を反映したもので、friends の動画の再生速度を変更した際に音割れが生じない下限が 0.6 倍速だったためである。また、偽陽性にて再生速度が低下してしまうことに対しカウンターバランスを取ることができると考え、笑いどころにて学習者が笑った場合、再生速度が 1 段階上がるようにした (図 5.3)。

## 第6章 本実験

ユーザの表情を検出し動画の再生速度を調節するシステム（以下、システム）の評価実験を行った。本章において、その調査内容および結果に関して述べる。

### 6.1 目的

本実験の目的は、第5章にて述べた動画の再生速度を調節するシステムの有用性を検証することである。このシステムは、外国語の動画を視聴している最中に、笑いどころにおけるユーザの表情を検出し、ユーザが笑っていれば再生速度を1段階上げ、笑っていなければ再生速度を1段階下げる。

### 6.2 機器

本実験には、ウェブカメラを内蔵したノート PC（Macbook Pro 13 inch, Big Sur バージョン 11.6, 2.8 GHz クアッドコア Intel Core i7, メモリ 16 GB）を用いた。PC を用いて、外国語の動画を参加者が視聴している最中、ウェブカメラの映像を用いて表情認識を行った（図 6.1, 図 6.2）。表情認識の処理速度は 30fps だった。実験参加者の顔が鮮明に見えるように、実験に用いた部屋の照明とウェブカメラの位置関係について、逆光になるのを避けた。PC ディスプレイに対し参加者の顔の距離は PC 操作時の標準的な距離である 40 cm から 60 cm であった [42]。PC ディスプレイと参加者の顔が正対するように PC を配置した。

### 6.3 実験参加者

本実験には、20 名（平均 22.6 歳、標準偏差 1.59 歳、16 名男性、4 名女性。P0, P1, ...P19 と以後表記）が参加した。20 名全員が英語を第二外国語として学んでいたため、英語を学習対象言語として選定した。実験参加者には実験参加前に、マスクを外す、および前髪を掻き分けるなど表情を見えやすくすることを指示した。また、普段の目の状態にて実験に参加するように指示した（裸眼 8 名、メガネ 5 名、コンタクト 7 名）。実験参加者の対象としては、CEFR スコア [62] にて B1 以上英語に習熟している者を対象とした。理由としては、多視聴といった動画を用いた外国語学習は、基本的な語彙を 2000 語から 3000 語ほど覚えている学習者を対象とした学習法であるということ [22,63]、および CEFR スコアにて B1 以上の学習者は、概ね 3000 語を超える語彙知識を有していること [20,21,64] から、動画を用いた外国語学習は



図 6.1: 本実験にて, ドラマを見ている様子.

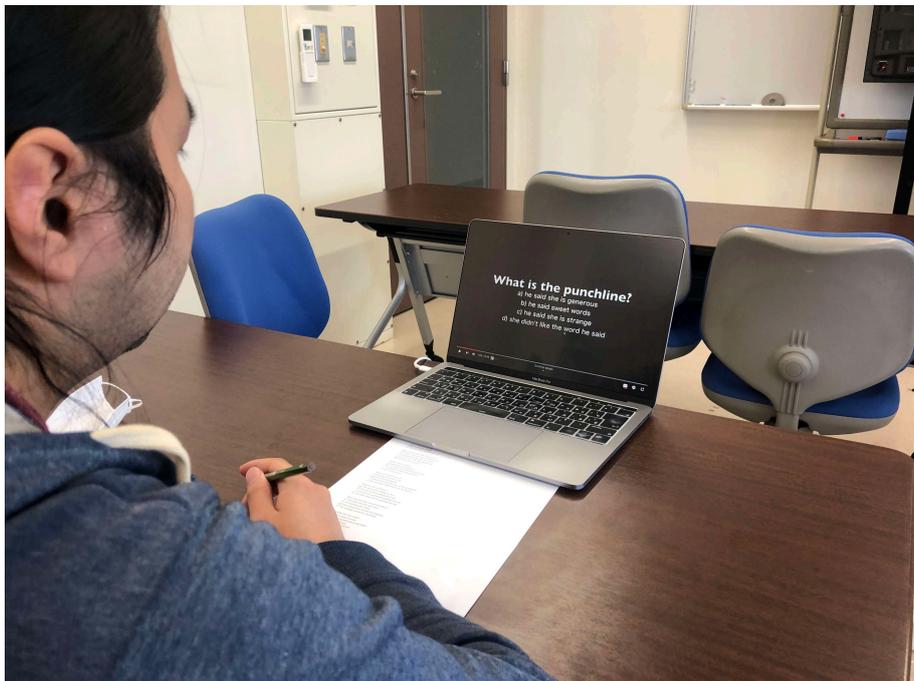


図 6.2: 本実験にて, 問題提示画面を見ている様子. 実験参加者は手元の問題兼回答用紙の選択肢のうち正答と考えられるものをマークする.

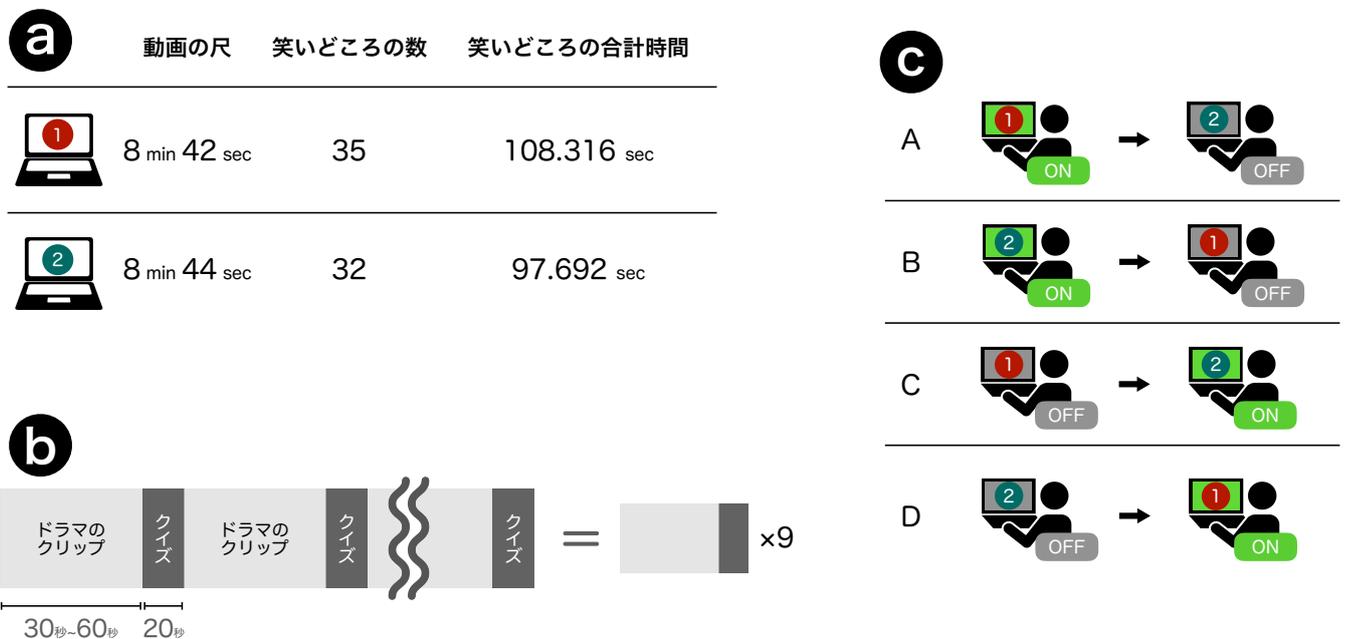


図 6.3: 統制の説明図. a) 動画の内容についての統制. b) 動画の構成. 9つのドラマのクリップおよびその内容に関するクイズから成る. c) グループのタスク統制.

CEFR スコアにて B1 以上の学習者が対象と考えられることが挙げられる. なお, CEFR スコアは TOEIC Listening&Reading Test のスコアに置き換えると, B1 は 550 から 785 (Listening スコア: 275 から 395, Reading スコア: 275 から 380) であると言われている [65].

## 6.4 実験設計

実験を行うにあたり, 動画の内容および動画内容に関するクイズ以外の構成をなるべく似せた動画を 2 本 (以後, 動画 1 および動画 2) 作成した (図 6.3 a, 図 6.3 b). それぞれの動画は 9 つのドラマのクリップおよびその動画内容に関する 4 択クイズ 9 問にて構成されており, それぞれの動画内に登場するドラマのクリップ, および動画 1 と動画 2 のドラマのクリップに前後関係は無かった. 動画 1 内にて発言される語数は 578 語, および動画 2 内にて発言される語数は 469 語だった. 動画に登場する語彙のテストもそれぞれの動画に対し 16-18 問作成した (動画 1 と動画 2 に共通して出現する語彙があったため, 群ごとに問題数の組み合わせが異なる). 語彙テストの問題の選定について, [66] に掲載されている語彙の出現頻度を

表 6.1: TOEIC の点数とグループ分けの内訳.

	TOEIC の点数					合計	平均 (標準偏差)
A	565	635	700	755	855	3510	702.00 (111.22)
B	550	690	755	760	815	3570	714.00 (101.82)
C	590	670	730	790	845	3625	725.00 (99.87)
D	P13	670	680	785	880	3015 <sup>†</sup>	753.75 <sup>†</sup> (98.94 <sup>†</sup> )

<sup>†</sup> P13 を除いた 4 人のスコアにて計算している.

参考に、話の流れを理解し笑いどころにおける発言を理解するために必要な語彙を選定した. 具体的にいうと, [19] を参考に, 出現頻度が少ないものから優先して 18 問を出題した. 動画内容に関する 4 択クイズについては, 第 3 章の通り実用英語検定, TOEIC Listening&Reading, および TOEFL iBT のリスニング解答時間<sup>1</sup>を参考にしつつ, 図 3.10 を考慮し, 1 問あたり 20 秒提示することにした. 複数の動画, 複数のクイズ, および複数の語彙テストを用いて実験参加者のパフォーマンスを測定することにより, 動画の内容によらない汎用的なシステムの評価が期待できる. 動画内容に関する 4 択クイズを出題することにより, 動画内容の把握具合を調べることができるとともに, 実験参加者の意識を動画に対し集中させることができる. 他, 語彙テストを動画の視聴前に行うことにより, 動画において語彙が分からず話の流れがわからなくなる可能性 (図 3.7) を減らすことができる. 同様に, 語彙テストを動画の視聴後に行うことにより, システムを用いることによる語彙やイディオムの定着効果も評価することができる. また, 実験参加者 20 人を A, B, C, および D の 4 グループに 5 人ずつ分けた (図 6.3 c). 全ての実験参加者に対し, システムを用いた際および用いない際のどちらの条件下においてもパフォーマンスを測定することにより, 語学力の個人差によらない汎用的なシステムの評価が期待できる. また, 4 グループに分けることにより, 動画 1 および 2 の見る順番に関する順序効果を相殺している. A から D までのグループ分けについて, 実験参加者は P13 を除き全員 TOEIC Listening&Reading Test<sup>2</sup>を受験したことがあったため, 筆者は参加者の TOEIC Listening&Reading Test のスコア (以下 TOEIC の点数) を元にグループ分けを行った<sup>3</sup>. なお, P13 は他に英語資格試験を受けたことがなかったが, P13 の大学 1 年次における英語の授業のクラス分けが最も初學者向けクラスだったこと, および P13 は英語の勉強を大学 1 年次以降特にしていないということから, 実験参加者の中において初學者に近い位置にいると判断し, 筆者はグループ分けを行った. TOEIC の点数およびグループ分けの内訳を表 6.1 に示す. 動画の内容については, 第 4 章で述べたように, 多視聴に用いられる動画はシットコムが良いとされている [8–10] ため, 再生する動画は YouTube の friends (シットコム

<sup>1</sup>実用英語検定のリスニング解答時間は 10 秒 (<https://www.eiken.or.jp/s-cbt/test/2020.html>). TOEIC のリスニング解答時間はおよそ 5 秒から 20 秒 (<https://eigo.plus/toEIC/toEIC-9>). TOEFL iBT のリスニング解答時間は 15 秒から 30 秒 [44].

<sup>2</sup><https://www.iibc-global.org/toEIC/test/lr.html>

<sup>3</sup>TOEIC は各テスト間の類似性が高いテストであるため, 英語の能力を測る指標として相応しいものである [67,68]

表 6.2: 動画前後の語彙テストや動画内容に関するテストの解答時間, ならびにシステムを用いた際の動画視聴時間.

	語彙テスト 1 または 2 の解答時間				設問の 先読み時間		システム使用時の 視聴時間
	事前		事後		クイズ 1	クイズ 2	
	テスト 1	テスト 2	テスト 1	テスト 2			
平均 (秒)	111.81	138.56	44.86	48.37	169.56	164.00	634.50
標準偏差 (秒)	31.27	56.57	13.02	19.51	86.79	92.38	44.62

の代表的作品) の動画クリップを用いた<sup>4</sup>. 動画を見せる際, システムを用いた条件 (実験条件) とシステムを用いない条件 (統制条件) のどちらで行っているかは実験参加者には知らせなかった (単盲検法にて実験を行った).

## 6.5 手順

統制により, A, B, C, および D はシステムを用いる順番および動画を見る順番が異なる (図 6.3 c). よって簡単のため, A グループを手本に説明する.

実験参加者はまず, 一通りの実験手順について説明を受けた後, 参加者自身のことについて調べるための実験前アンケートに回答した (付録 B.1). その後, 参加者は動画 1 に登場する語彙のテスト (以下語彙テスト 1) に解答した (付録 B.2). 参加者が語彙テスト 1 に答えたのを確認後, 語彙テスト 1 の正解を教えた. これにより, 実験参加者は動画内容に必要な語彙を知ることができるため, 動画において語彙が分からず話の流れがわからなくなる可能性 (図 3.7) を減らすことができる. 次に, 動画 1 の内容に関する 4 択クイズ (以下動画内容クイズ 1) について, 9 問全ての問題文および選択肢を動画視聴前に読んでもらった. これにより, 実験参加者は問題文および選択肢に分からない語彙がないかを確認めつつ, 解答時間内に問題文や選択肢を読み切ることが充分容易になる. よって, 語彙が分からず動画内容クイズ 1 に間違えたり, 解答時間内に問題文や選択肢を読みきれず不正解になる可能性を減らすことができる. その後, 実験参加者は動画 1 をシステムを使用した状態にて視聴する. 視聴中, ドラマのクリップが流れている場面において, 笑いどころにおける参加者の表情に従ってシステムが作動し動画の再生速度を調節する (図 6.1). また, 問題文が提示された場面において, 参加者は 20 秒の解答時間内に問題を解く (図 6.2). 動画を見終わったのち, 参加者は語彙テスト 1 に再度解答し, 動画 1 をシステムを用いて視聴した後の語彙の定着度を測った. その後, 参加者は作業負荷を測定するために NASA-TLX [69] に回答した. また, 参加者は 5 分間以上の休憩ののち, 先述した実験条件下における手順と同様の手順を対照条件下にて行った. この条件下においては, 参加者は動画 1 の代わりに動画 2 を, システムを使用せずに視聴しする. 動画内容に関するクイズおよび語彙テストについても, 参加者は動画 2 の内容に即したもの (動画内容クイズ 2 および語彙テスト 2) に答えた (付録 B.3). 最後に, 参加者は SUS [70] および実験後アンケートに回答した. 実験後アンケートでは実験中に感じた

<sup>4</sup><https://www.youtube.com/watch?v=nvzkHGndt fk>

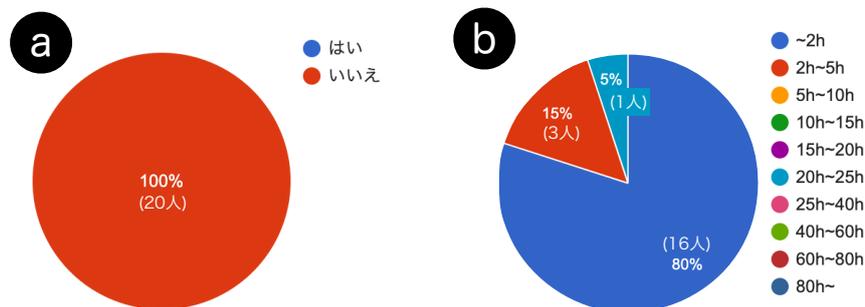


図 6.4: 事前アンケートの結果の一部. a) 英語はあなたの母語ですか. b) 日常生活で英語はどれくらい使いますか (時間/週)

ことやシステムを用いた感想について聞いた. 実験の合計時間は1人当たり平均64分だった. 語彙テストの解答時間, 動画内容に関するクイズの設問を事前に見た時間 (設問の先読み時間), システム使用時の動画視聴時間を表 6.2 に示す. システムを用いていない際の動画時間は8分43秒 (523秒) なので, およそ1.213倍視聴時間が伸びている.

実験の際の留意事項としては, 録画をしていると参加者が緊張し顔がこわばる可能性がある [71] ため, 実験を通して録画をしていないことを参加者に伝えた. また, 参加者が緊張しないように, 参加者が動画を見ている最中, 実験者は部屋から出た. その他, 動画内容の先読みにより動画内容クイズに影響が出ないように, 語彙問題および動画内容クイズに答える際, 動画内容を極力思い浮かべないように参加者に伝えた.

## 6.6 収集データ

実験にて収集し, 解析に用いたデータは以下の通りである.

- システムを用いて動画を見た際の再生速度の変更記録
- 実験中の動画の内容に関する4択クイズに対する解答
- 実験前アンケート (付録 B.1)
- 実験中の語彙テスト (付録 B.2 および付録 B.3)
- 実験後アンケート (付録 B.4)

## 6.7 結果

以下では, 本実験にて取得したそれぞれのデータの解析を行った結果を示す.

表 6.3: 語彙テストの点数. テストの点数の上限値が 16 点から 18 点と異なるため, 各セルに上限値を書いている.

	平均値 (標準偏差) / 上限値				中央値/上限値			
	動画 1		動画 2		動画 1		動画 2	
	視聴前	視聴後	視聴前	視聴後	視聴前	視聴後	視聴前	視聴後
A	15 (1.22) /18	18 (0.00) /18	9.6 (2.61) /16	15.8 (0.45) /16	15/18	18/18	10/16	16/16
B	13.8 (1.30) /17	17 (0.00) /17	11.6 (2.70) /17	16.8 (0.45) /17	14/17	17/17	11/17	17/17
C	14.2 (1.64) /17	17 (0.00) /17	10.6 (1.67) /17	17 (0.00) /17	14/17	17/17	11/17	17/17
D	15.6 (0.89) /18	17.6 (0.55) /18	9.2 (2.95) /16	16 (0.00) /16	15/18	18/18	11/16	16/16

### 6.7.1 実験前アンケート

本実験前のアンケートについて, 主要なものを以下に示す.

- 英語はあなたの母語ですか (図 6.4 a).
- 日常生活で英語はどれくらい使いますか (時間/週) (図 6.4 b).

実験参加者全員が英語を第二外国語として学んでおり, 母語話者はいなかった. また, 日常生活にて英語を用いる時間について, 週に 2 時間以下という意見が大多数を占めていた.

### 6.7.2 実験中アンケート: 語彙テスト

動画 1 および動画 2 の視聴前後に, 実験参加者は語彙テストに答えた. 各群の語彙テストの点数を表 6.3 に示す. システムを用いた際も用いていない際も, 動画視聴後の語彙テストの点数は全員ほぼ満点であり, システムを用いることによる語彙の定着度への差は見られなかった.

### 6.7.3 システムを用いて動画を見た際の再生速度の変更記録

システムを用いて動画を見た際, 再生速度の変更記録を取得した. 機器の不具合により, A グループの P17, B グループの P19, および D グループの P18 の再生速度の変更記録は取得できなかった. 残りの参加者の動画視聴時の再生速度の変更記録について, グループ毎の再生速度の変化を図 6.6 に示す. グループ毎に見ても, 再生速度の変更記録に顕著な特徴は見られなかった.

実験参加者が動画内のどの笑いどころにおいて笑っており, どの笑いどころにおいて笑わなかったかをプロットした図が図 6.5 である. この図を見ると, ほとんどの笑いどころにおいて笑った者と笑っていない者がおり, 笑いどころにおいて笑うか否かは個人差があることが分かる.

また, TOEIC の点数別に再生速度の変化を示した図が図 6.7 である. この図を見ると, 初めから再生速度が低下し続け, 下限である 0.6 倍速にて最後まで一定の者が 3 名いること, お

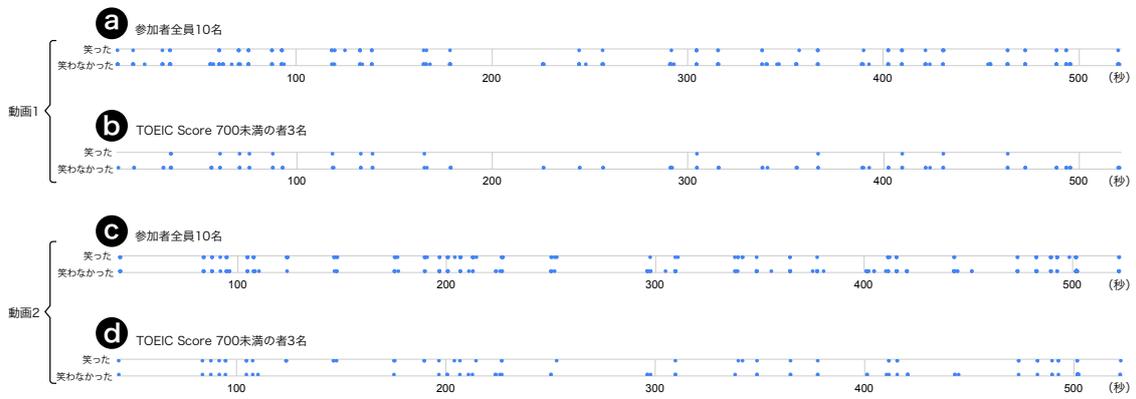


図 6.5: 動画の経過時間に対し、笑った点および笑っていない点の時刻をプロットした図. a) 動画1を見た者全員分のもの. b) 動画1を見た者のうち、TOEICの点数が700未満の参加者のもの. c) 動画2を見た者全員分のもの. d) 動画2を見た者のうち、TOEICの点数が700未満の参加者のもの.

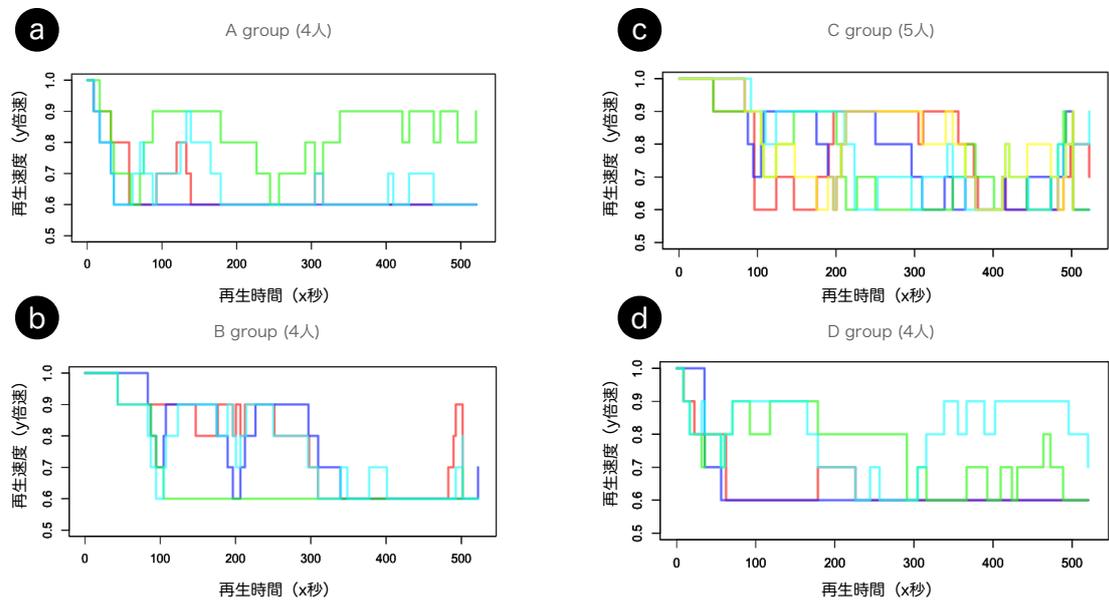


図 6.6: グループ毎の再生速度の変化. a) Aグループ. b) Bグループ. c) Cグループ. d) Dグループ.

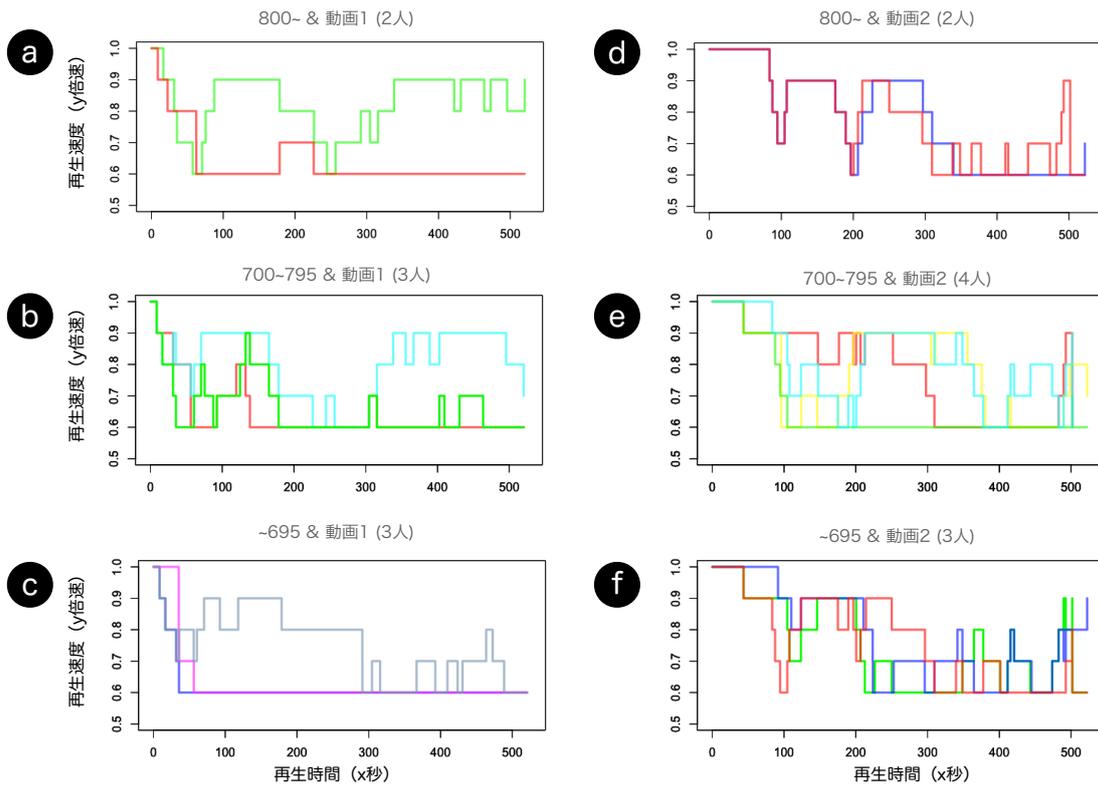


図 6.7: TOEIC のスコア毎の再生速度の変化. a) TOEIC の点数 800 以上で動画 1 を見た群. b) TOEIC の点数が 700 以上 795 以下で動画 1 を見た群. c) TOEIC の点数が 695 以下で動画 1 を見た群. d) TOEIC の点数 800 以上で動画 2 を見た群. e) TOEIC の点数が 700 以上 795 以下で動画 2 を見た群. f) TOEIC の点数が 695 以下で動画 2 を見た群.

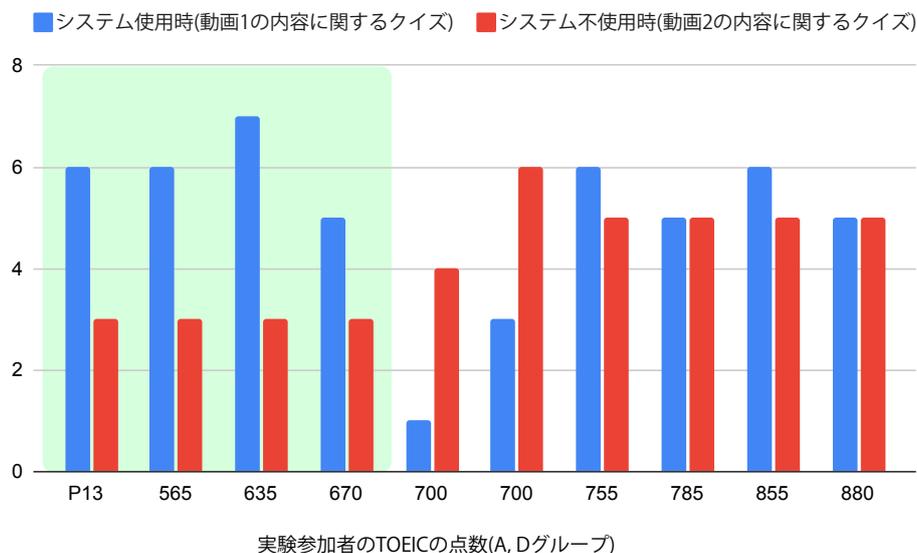


図 6.8: A グループと D グループにおける動画内容に関するクイズの正答数と参加者の TOEIC の点数の関係. A グループと D グループは動画 1 をシステムを用いて視聴し, 動画 2 をそのまま視聴した.

よび, 動画の種類によってある程度再生速度の変化がパターン化されていることが分かる. この図からも図 6.5 と同じく, 語学力や笑いのツボといった実験参加者の個人差に結果が影響されていることが示唆される. しかし, この結果から, 動画内容が再生速度の変化に与える影響も存在するように見える.

#### 6.7.4 動画内容に関するクイズ

動画内容に関するクイズを動画内に配置し, 参加者に解答させた. その際のクイズの正答数と参加者の TOEIC の点数の関係を図 6.8, 図 6.9, 図 6.10, および図 6.11 に示す. 図 6.8 および図 6.9 はシステムを用いた/用いない動画の組み合わせが一致するグループ別に分けたグ

表 6.4: 動画内容に関するクイズの正答数と参加者の TOEIC の点数の相関関係の表.

システムを使用したか	視聴した動画/グループ	相関係数 (Spearman)	無相関検定 (Spearman) の有意水準
システム使用	動画 1/ (A, D)	-0.22	0.58
	動画 2/ (B, C)	-0.11	0.77
システム不使用	動画 1/ (B, C)	0.57	0.08
	動画 2/ (A, D)	0.72	0.03

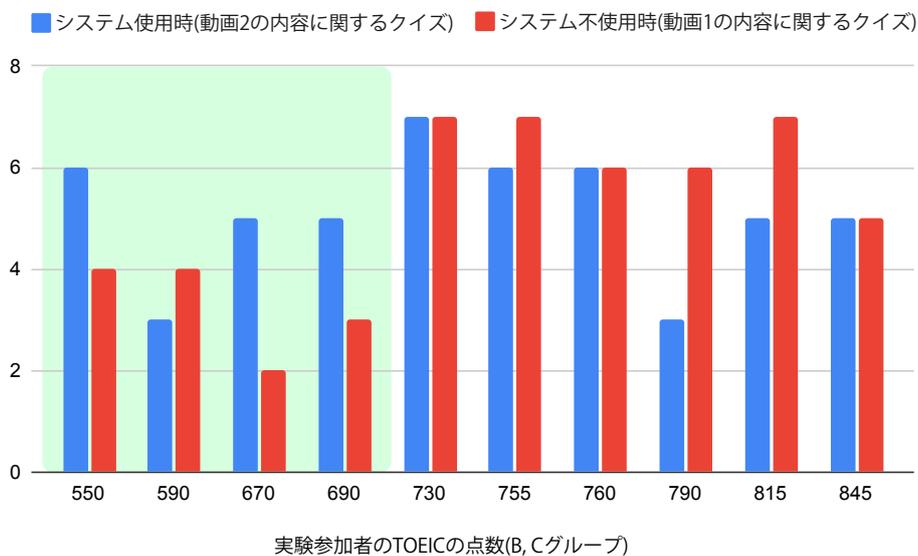


図 6.9: B グループと C グループにおける動画内容に関するクイズの正答数と参加者の TOEIC の点数の関係. B グループと C グループは動画 2 をシステムを用いて視聴し, 動画 1 をそのまま視聴した.

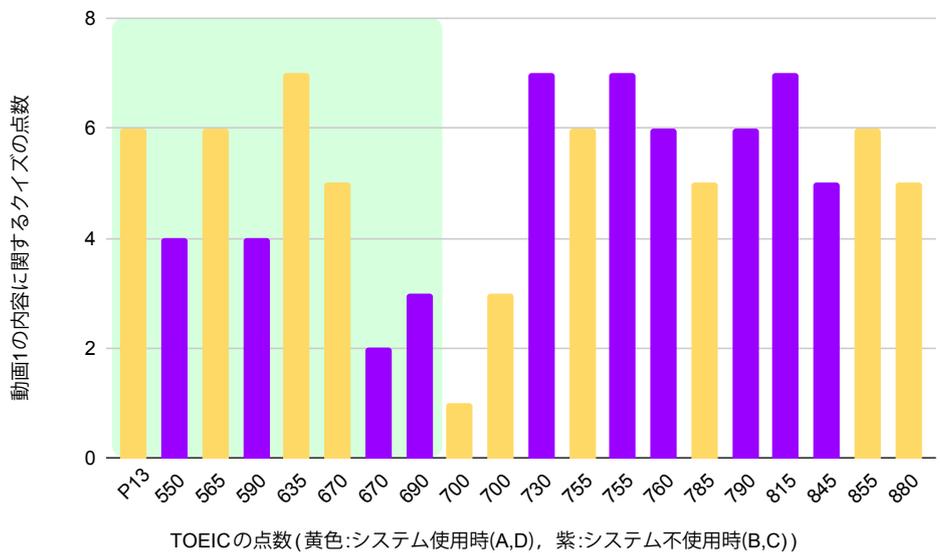


図 6.10: 動画 1 の内容に関するクイズの正答数と参加者の TOEIC の点数の関係. A グループと D グループはシステムを用いて視聴し, B グループと C グループはそのまま視聴した.

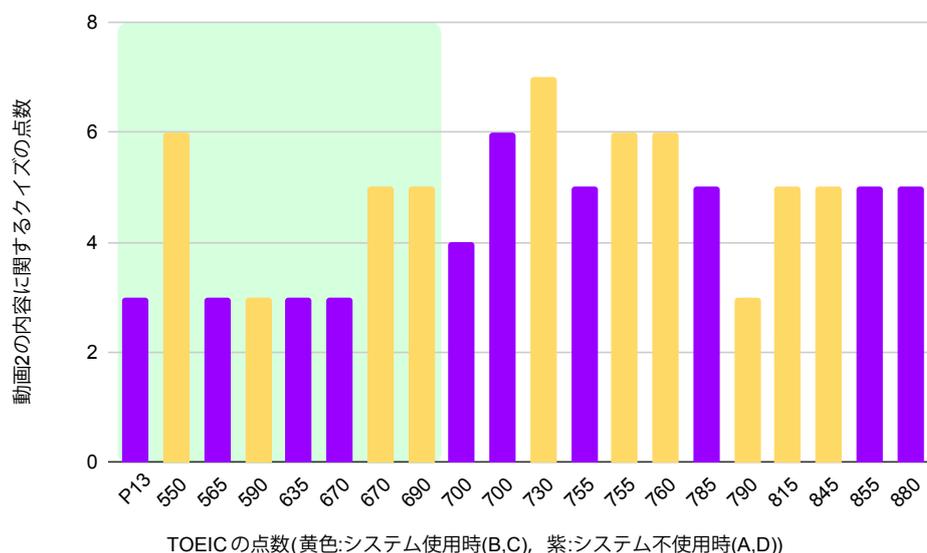


図 6.11: 動画 2 の内容に関するクイズの正答数と参加者の TOEIC の点数の関係. B グループと C グループはシステムを用いて視聴し, A グループと D グループはそのまま視聴した.

表 6.5: 動画内容に関するクイズの正答数の二群間検定の結果.

TOEIC の点数 による区分け		中央値		平均値 (標準偏差)		有意確率	効果量	効果量の 95%信頼区間	
		システム有	システム無	システム有	システム無			下限	上限
700 未満	クイズ 1	6.0	3.5	6.00 (0.82)	3.25 (0.96)	p<.005	2.69 (大)	0.61	4.76
	クイズ 2	5.0	3.0	4.75 (1.26)	3.0 (0.0)	p<.05	1.71 (大)	-0.05	3.47
700 以上	クイズ 1	5.0	6.5	4.33 (1.97)	6.33 (0.82)	p<.05	1.23 (大)	0.07	2.52
	クイズ 2	5.5	5.0	5.33 (1.37)	5.0 (0.63)	0.39	0.29 (小)	-0.90	1.48
全体	クイズ 1	5.5	5.5	5.0 (1.76)	5.1 (1.79)	0.85	0.05 (無視可能)	-0.85	0.95
	クイズ 2	5.0	4.5	5.1 (1.29)	4.2 (1.14)	0.10	0.71 (中)	0.22	1.64

ラフであり、図 6.10 および図 6.11 は動画別に分けたグラフである。それぞれの図を見ると、システムを用いていない条件下（赤色および紫色）において、TOEIC の点数とクイズの正答数には正の相関関係があるように見える。また、システムを用いている条件下（青色および黄色）において、TOEIC の点数とクイズの正答数には相関関係がないように見える。よって、クイズの正答数と TOEIC の点数の相関検定を行った表 6.4。さらにいうと、TOEIC の点数が比較的低い層を見ると、動画 1 を用いたシステムを用いた群の方が用いていない群よりも点を取っているように見える。よって、システムを用いた群と用いていない群においてサンプルサイズが一致して統計的処理を行いやすいため、TOEIC の点数が比較的低い層を 700 未満（緑の範囲）と設定し、TOEIC の点数が 700 未満のサンプル、700 以上のサンプル、および全てのサンプルについて、クイズの正答数の二群間の比較検定を行った表 6.5。線形性を仮定しないため、相関分析および無相関検定には Spearman の相関係数を用いた。正規性の確認には Shapiro-Wilk 検定および Kolmogorov-Smirnov 検定、等分散性の確認には F 検定、正規性および等分散性を仮定した二群間の比較検定には対応のない t 検定、正規性のない二群間の比較検定には Wilcoxon の順位和検定、および効果量の算出には Hedges' g を用いた。効果量の解釈には [53] を用い、「大（大きい）」「中」「小（小さい）」「無視可能」の 4 段階にて評価した。

表 6.4 より、システムを用いなかった群はクイズの正答数と TOEIC の点数に有意に正の相関関係があることを示したが、システムを用いた群には相関関係は見られなかった。

表 6.5 より、TOEIC の点数が 700 未満のサンプルにおけるシステム使用群と不使用群の比較において、システムを用いた際、動画内容に関するクイズの点数に有意な上昇が見られ、効果量も大きかった。また、TOEIC の点数が 700 以上のサンプルにおけるシステム使用群と不使用群の比較において、システムを用いた際、動画内容に関するクイズの点数には有意な下降が見られたか、もしくは有意差が見られなかった。全体的にシステム使用群と不使用群の比較を行ったところ、有意差は見られなかった。

### 6.7.5 実験後アンケート

本実験後のアンケートについて、主要なものを以下に示す。

- デバイスを用いた時に動画コンテンツを楽しむ事はできましたか（図 6.12 a）。
- デバイスを用いなかった時に動画コンテンツを楽しむ事はできましたか（図 6.12 b）。
- デバイスを用いた時の動画の速度調整は適切でしたか（図 6.12 c）。
- 表情検出から自動で再生速度が変わるシステムを使用した際、自分で速度を調整した方がいいと感じた場面がありましたか（図 6.12 d）。
- 上記の質問で「5, 6, 7」を選んだ場合（自分で速度を調整した方がいいと感じた場面があった場合）理由を書いてください。
- よりよく理解できた動画は動画 1 と動画 2 のどちらですか（図 6.13 a）。

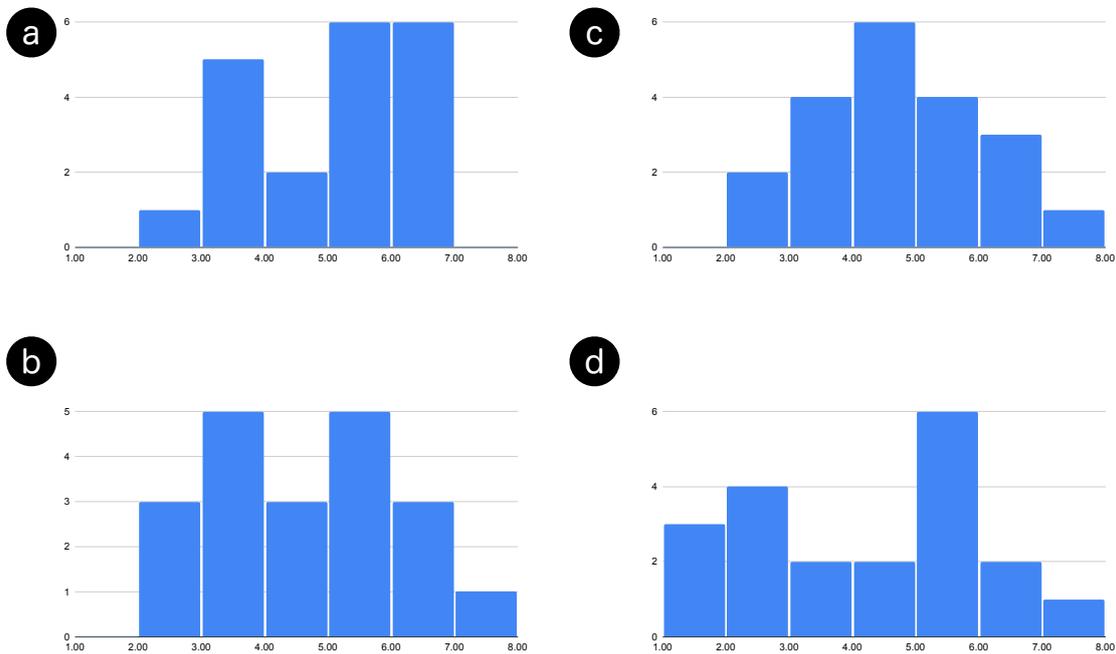


図 6.12: 事後アンケートの結果の一部. a) デバイスを用いた時に動画コンテンツを楽しむ事はできましたか. b) デバイスを用いなかった時に動画コンテンツを楽しむ事はできましたか. c) デバイスを用いた時の動画の速度調整は適切でしたか. d) 表情検出から自動で再生速度が変わるシステムを使用した際, 自分で速度を調整した方がいいと感じた場面がありましたか.

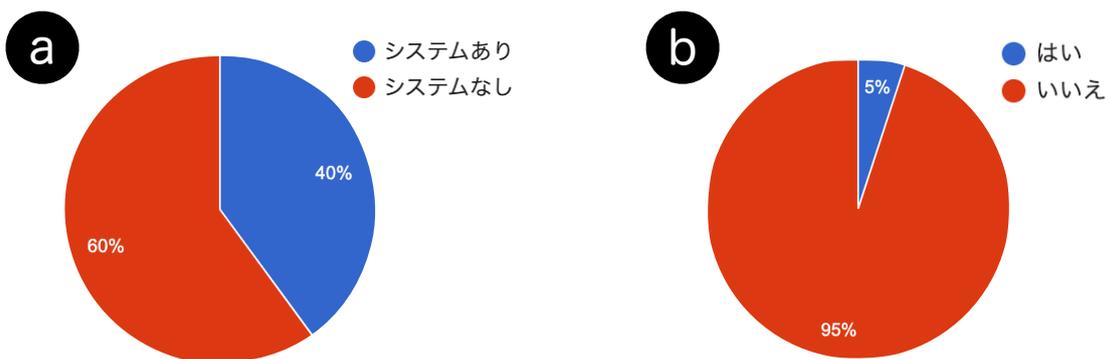


図 6.13: 事後アンケートの結果の一部. a) よりよく理解できたのは動画1ですか動画2ですか (ただし, システム使用時/不使用時という可視化の仕方にした). b) 今回の動画コンテンツ (friends というドラマ) は以前に試聴したことがありますか.

表 6.6: NASA-TLX のスコアの表.

		知的・知覚的要求	身体的要求	タイムプレッシャー	作業成績	努力	フラストレーション	総合得点
システムを用いた時	平均	74.00	18.00	56.75	59.75	62.50	46.00	64.43
	標準偏差	18.54	17.43	27.45	19.50	22.39	22.16	11.64
システムを用いなかった時	平均	80.25	23.25	68.25	59.50	70.75	47.50	71.05
	標準偏差	14.19	23.52	28.25	21.51	20.98	26.18	15.29

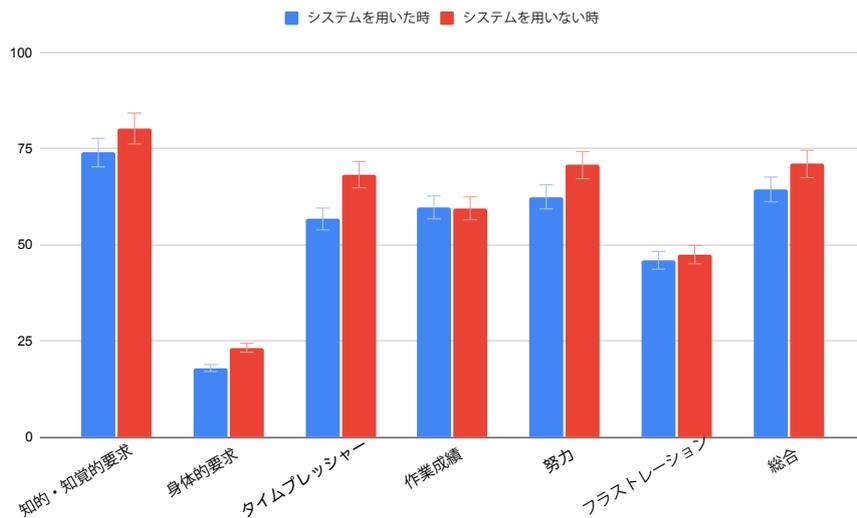


図 6.14: NASA-TLX のスコアの平均値および 10% 標準誤差.

- 今回の動画コンテンツ (friends というドラマ) は以前に試聴したことがありますか (図 6.13 b) .
- 動画内で最後までわからなかった単語があれば書いてください.
- その他気になること, 気づいたことはありましたか.

動画内で最後までわからなかった単語について, generous という回答が 1 つあった.

### 6.7.6 NASA-TLX による作業負荷の指標

システムを用いて動画を見た際, およびシステムを用いずに動画を見た際に NASA-TLX を用いて実験参加者の作業負荷を測った. それぞれの NASA-TLX のスコアを表 6.6 および図 6.14 に示す. システムを用いた時の方が, 作業成績以外の全項目において有意に負荷が小さかった.

### 6.7.7 SUSによるユーザビリティの評価

実験後、SUS [70]によるシステムのユーザビリティ評価を行った。その結果、20人分のスコアは82.5, 67.5, 90, 75, 60, 60, 85, 85, 80, 60, 80, 87.5, 85, 45, 85, 90, 52.5, 47.5, 72.5, および85だった。これらの平均値は73.75で、標準偏差は14.65614655, 中央値は80だった。[72]によると、SUSスコアが示すユーザビリティの評価はGoodおよびAcceptableである。

SUSスコアの分布の特徴について、80点前後を申告したものと、50点前後を申告したものに二極化されているように見える。

### 6.7.8 システムに対する意見

「表情検出から自動で再生速度が変わるシステムを使用した際、自分で速度を調整した方がいいと感じた場面がありましたか」(図 6.12 d)という質問に対し、リッカート尺度にて「5, 6, 7 (感じた)」を選んだものに対し、「上記の質問で「5, 6, 7」を選んだ場合(自分で速度を調整した方がいいと感じた場面があった場合)理由を書いてください」という追加質問をした。それに対し、次のような回答をもらった。

- ゆっくりで丁度いい箇所もあったが、ゆっくりすぎてもう少しだけスピードを変えたいと感じた箇所もあった。
- 動画を見て問題を解いているときに、分からないの顔が反映されたのか、次の動画で遅くなった気がするから。
- 全体の速度調整は自分で行えてもいいかもしれない。
- 遅く感じるのでフラストレーションが溜まった部分もあった。
- 遅いと感じることが何度かあったことと、動画中に速度が変わっている感じがあまりしなかったため。
- 確かに通常のスピードで聞き取れなかったが、流石に遅すぎて英語の発音は全く関係なく字幕だけ見ていたのもう少し早くしたいと思ったため。
- かなり遅く感じたため。
- 速度が遅すぎて逆に分かりづらいシーンがあったため。
- 文章の意味がわからない時(スラング的な)。

これらの回答から、動画の再生速度が遅いと感じた場合、参加者はフラストレーションを感じる事が分かった。

また、「その他気になること、気づいたことはありましたか」という質問に対する解答としては次のような回答をもらった。

- システムがコメディにとってもあっている。その他のジャンルにも適用できたら有効かどうかはわからないが面白そう。
- ゆっくりになり過ぎていることがある気がした。
- 英語が難しかった。後半速度が落ちたように感じた。
- 1本目の動画でリスニングを久しぶりに行ったので2本目ではだいぶ慣れがあったと思った。デバイスを用いた場合に、発音が遅く、読むのには役に立ったが発音自体は全く聞いていなかった。1本目の遅いスピードに慣れて2本目の1つの動画が短く感じてこれで終わり？と思った。(本実験参加者はグループAに属していた)
- どこでゆっくりになったのかあんまり分からなかった(なんとなくゆっくりになっている気はした)。分かってなくてもつられて笑っちゃう時があった。
- 1の方が内容が平易だと感じた。
- もしかしたら、ループするとかの方がいいのかもしれないです。
- 動画内のストーリーが突飛なもので、理解が追いつかないことがしばしばあった。
- 内容的に動画2の方が難しい
- 会話のスピードが遅くなることで聞き取りやすくなる場面が何度かあった。ゆっくりすぎてもう少しだけ早くして欲しい箇所もあった。
- 動画再生中は字幕に目が入っていた。
- 分からなくても顔にでるタイプではないため、速度が遅くなることが少なかった。また速度が遅くなるとスピーディさがなくなりコンテンツを楽しむという点においては、速度調整がない方が楽しめた。また速度調整が遅くなった時は字幕を読む時間が増え内容理解を促進するが、分からない単語であるとうとうしようもない。
- 再生速度がスローになった時に、独特の低い声がコンテンツの面白さを半減させていたように思う。感嘆詞や相槌などはスローだと獣の雄叫びみたいになって聞き取れないことがあった。なので本研究から外れるかもしれないが、スローになった時の声がもう少し自然になるとコンテンツの満足度が大幅に向上すると感じた。
- システムを使ったことに気づけなかった。自然に速度が変化していたということだと思ふ。
- 再生速度の変化により会話の「間」が変化したため、会話の中で生まれた沈黙なのか、速度変化によって生まれた静止なのかを見分けることが困難だった。
- ゆっくりしている方が聞きやすかった。

この結果から、動画速度が自分の思った通りでないと感じた参加者が一定数いる一方で、速度調節が行われたことに気づかなかったり、速度調節が心地よいと感じた参加者もいることが分かった。また、タイムストレッチが実装されている YouTube をプラットフォームとして選んだものの、ピッチ変化が気になると回答した参加者もいることが分かった。字幕を理解の手助けとして用いている参加者も多いことが分かった。

## 6.8 考察

語彙テストに満点が出たため、システムを用いた際の語彙の勉強効果はうまく測定できなかった。動画視聴の直前および直後のタイミングにて測定しているため、間隔が 10 分程度しか空かなかったことが原因と考えられる。もっと長い期間で実験をできていないため、詳しい追跡が可能な実験参加者を募集することにより語彙の長期記憶に与える影響を調べることができる。と考える。

また、TOEIC の点数別に再生速度の変化を示した図 6.7 を見ると、初めから再生速度が低下し続け、下限である 0.6 倍速にて最後まで一定の者が一定数いること、および、動画の種類によってある程度再生速度の変化がパターン化されていることが分かる。実験参加者の厳密な語学力や笑いのツボに結果が影響されている可能性はあるが、動画内容が再生速度の変化に与える影響は大きいように見える。

参加者の TOEIC の点数と動画内容に関するクイズの点数の比較については、検定内において、参加者という変数と、システムを使ったか否かという変数が異なるため厳密には比較ができていない。しかし、サンプル数を増やせば参加者の違いから来る結果への影響は少なくなるはずである。

実験結果からわかることとして、動画内容クイズの結果より、本システムの実装を実験動画に適用した場合、TOEIC の点数がおよそ 700 未満 (CEFR スコア B1 下位以下) の学習者に特に有用であったと言える。しかし、それ以上の TOEIC の点数を取得している学習者にとってはシステムを使った効果があまり見られなかったため、本システムとは異なる支援方法を考える必要がある。

動画内で最後までわからなかった単語について、generous という回答が 1 つあった。しかし、それ以外は全員無回答だった。したがって、実験参加者は動画視聴中に出現する語彙知識に関する問題に直面しなかったと考えられる。

SUS スコアについて、スコアが Acceptable や Good と言えるので、本システムは良いユーザビリティを有すると考えられる。しかし、アンケート項目の「よりよく理解できた動画は動画 1 と動画 2 のどちらですか」を見ると、システムを用いていない時に視聴した動画の方に参加者は票を入れているため、普通の速度にて見る方が参加者の自覚的にはより理解できたと解釈されていると考えられる。

## 第7章 議論

動画コンテンツの話自体が学習者にとって笑えないものだった場合、シリアスな動画コンテンツを視聴しており笑いの場面がない場合、あるいは非言語的な背景知識が無く笑えなかった場合におけるシステムの誤作動が問題として残る。よって、より適切に理解度を計測する手法 (e.g. 脳波測定, 瞳孔径測定) を探索する予定である。また、録音笑いの検出についても、機械学習モデルを用いたが、笑いどころ推定の誤認識も考え、録音笑いや観客の笑い声がなかった場合においても適切に笑いどころを測定する手法についても探索する予定である。

さらに、本実験においては、ブラウザの操作などの既存の動画速度変更手法との比較を行っていない。引き続き、実験によって既存手法との差を調べる必要がある。

また、シットコムは通常、隣接するエピソード間の繋がりががあるため、本実験にて提示したように、無関係のクリップを集めたものだと、ドラマとして面白くない可能性がある。システムを用いたときの楽しさについてはシステムを用いた際および用いていない際にて差がなかったため、繋がりのあるエピソードの統制条件にて実験する価値がある。

ならびに、実験において用いた動画の数にも制限がある。本実験においては、2本の動画を用いたに過ぎないのもっと多様な動画を使う必要がある。実験参加者数をより増やせば、厳密にこのシステムの有効な層がより具体化される。

さらに、巻き戻し、動画を繰り返し見る学習法も試すと良いと第4節にて数々の国内会議参加者から言われた。巻き戻しについては、繰り返し見ることにより動画視聴のテンポが悪くなり、多視聴に必要である娯楽性が失われる可能性が考えられるので今回の実装には用いなかったが、今後の実験にて実装および検証する。

また、グラフの可視化による様子から、今回は TOEIC の点数 700 にて実験参加者のレベル帯を区切った。しかし、実験参加者数を増やせばシステムの効果が得られる層および動画の難易度の関係がより判明すると考えられるため、増やすことを考えている。

また、動画1については TOEIC の点数 700 にて実験参加者のレベル帯を区切るのが適切であったと考えられる (図 6.10) が、動画2については TOEIC の点数 700 より上の層においてもシステムを用いた方がスコアが高くなっている (図 6.11)。アンケート結果において、動画1よりも動画2の方が内容的に難しかったという意見が複数見られたことから、動画2についてはより英語力の高い層までシステムが有用だった可能性がある。

また、動画内容について、個人の背景知識が内容の理解に影響した可能性がある [73]。これについては、厳密に統制を図る事は難しい。

## 第8章 今後の展望

この研究は将来的に、英語学習や語学学習以外にも、MOOCsの動画を教材に選べば、逆に視聴時間を短縮し勉強効率を上げるなどの実装可能性もありうる。また、筑波大学という外国語の学習選択肢および学習者が多い環境を生かし [74,75]、ロシア語など、英語以外の外国語を用いた実験を行い、今回の本実験で得られた結果と同様の結果が現れるか検証することを考えている。

また、ほかにいくつかの動画を用いて実験を行うことも考えている。そうすることで、教材の難易度を加味し、学習者の外国語の習熟度に合わせた適切な難易度の外国語学習教材を勧めることができる。

## 第9章 おわりに

本稿では、外国語の動画コンテンツを用いて外国語学習を行う際に、学習者が動画コンテンツの再生速度についていけない場合に備え、学習者の理解度に合わせて再生速度自動調節機能を備えたシステムを提案した。学習者の理解度については、予備実験の結果およびデモ発表におけるフィードバックに基づき、表情および笑いに関するコンテンツの文脈から判断することにした。

評価実験において、本システムは全ての実験参加者の作業負荷を高めず、英語が中級レベル(CEFR スコア: およそ～B1 下位)の参加者において、内容に対する理解の向上が有意に見られた。すなわち、本システムを用いることにより、多聴学習や多視聴に代表される外国語の動画を視聴する学習方法において、個人のリスニング能力に最適化された速度にて動画を視聴することができることが分かった。このシステムを用いて学習することにより、ビデオ教材の難易度を調整することができるため、多視聴や多聴学習を行う際に用いることができる学習教材の選択肢が広がる可能性が示された。

## 謝辞

本研究を進めるにあたり、志築文太郎先生、高橋伸先生には多くのご意見とご指導を頂きました。心から感謝いたします。特に志築文太郎先生には、研究の進め方、論文執筆をはじめとして研究に関することを基礎からご指導いただきました。また、研究に関する相談にも乗っていただき、多くのご助言を頂きました。奨学金書類や院試への推薦書への執筆も行っていただきました。重ねて感謝いたします。インタラクティブプログラミング研究室の同期、先輩方には様々な面において多くのご助言を頂きました。特に WAVE チームの皆様にはチームゼミにおけるご意見や論文の添削といった研究に関するだけでなく、学部卒業後の進路の相談といった人生についても相談に乗っていただき多くのご意見、アドバイスを頂きました。深く感謝いたします。中でも、野崎陽奈子氏にはメンターとして研究方針、実験の方法、論文執筆など本研究における重要な部分において多くのアドバイスを頂き大変お世話になりました。ここに厚くお礼申し上げます。また、研究生活に限らず、多くの面において支えて頂いた友人、大学生活にてお世話になった先生方、皆様に感謝申し上げます。特に、学内外の人文系および文化系に所属する先生方および友人達には課外活動、授業、海外研修、および科目履修プログラムを通じ目をかけていただき、仲良くしていただき、情報系知識を超えた広い知見を与えてくださったことに非常に感謝しております。また、実験に参加していただいた皆様の貢献は本研究の遂行に不可欠なものでした。深く感謝いたします。最後に、普段から私を支えてくださった家族に深く感謝いたします。

## 参考文献

- [1] Willy Renandya and George Jacobs. *Extensive Reading and Listening in the L2 Classroom*, chapter 8, pp. 97–110. Springer, 2016.
- [2] Ahmed Masrai. Can L2 Phonological Vocabulary Knowledge And Listening Comprehension Be Developed Through Extensive Movie Viewing? The Case Of Arab EFL Learners. *International Journal of Listening*, Vol. 34, No. 1, pp. 54–69, 2020.
- [3] Willy A Renandya and Thomas SC Farrell. ‘Teacher, the tape is too fast!’ Extensive Listening in ELT. *ELT journal*, Vol. 65, No. 1, pp. 52–59, 2011.
- [4] Willy A Renandya. The Tape Is Too Fast. *Modern English Teacher*, Vol. 21, No. 3, pp. 5–9, 2012.
- [5] Wang Li and Willy A Renandya. Effective Approaches to Teaching Listening: Chinese Efl Teachers’ Perspectives. *The Journal of Asia TEFL*, Vol. 9, No. 4, pp. 79–111, 2012.
- [6] Anna CS Chang. *Teaching L2 Listening: In and Outside the Classroom*, pp. 111–125. Springer, 2016.
- [7] Betül Bal-Gezegin. An Investigation of Using Video vs. Audio for Teaching Vocabulary. *Procedia - Social and Behavioral Sciences*, Vol. 143, No. 3, pp. 450–457, 2014.
- [8] Webb Stuart. Extensive Viewing: Language Learning through Watching Television. *Language Learning Beyond the Classroom*, pp. 159–168, 2015.
- [9] Mehmet Özgen and Nazlı Gunduz. Authentic Captioned Sitcom as Listening Comprehension Material in English Language Teaching. *ELT Research Journal*, Vol. 9, No. 2, pp. 167–193, 2020.
- [10] Paulo Quaglio. *Television Dialogue: The sitcom Friends vs. natural conversation*, Vol. 36. John Benjamins Publishing, 2009.
- [11] Natsuki Fukunaga. “Those Anime Students”: Foreign Language Literacy Development Through Japanese Popular Culture. *Journal of Adolescent and Adult Literacy*, Vol. 50, No. 3, pp. 206–222, 2006.

- [12] Hung chun Wang and Cheryl Wei yu Chen. Learning English From Youtubers: English L2 Learners' Self-regulated Language Learning on Youtube. *Innovation in Language Learning and Teaching*, Vol. 14, No. 4, pp. 333–346, 2020.
- [13] Elke Peters, Eva Heynen, and Eva Puimège. Learning Vocabulary Through Audiovisual Input: the Differential Effect of L1 Subtitles and Captions. *System*, Vol. 63, pp. 134–148, 2016.
- [14] Chihiro Fujimori. The Effects Of An Extensive Reading Program On Reading And Listening Comprehension Among Senior High School Students. *The Bulletin of the Kanto-koshin-etsu English Language Education Society*, Vol. 21, pp. 13–24, 2007.
- [15] Yanxue Feng and Stuart Webb. Learning Vocabulary Through Reading, Listening, and Viewing: Which Mode of Input Is Most Effective? *Studies in Second Language Acquisition*, Vol. 42, No. 3, pp. 499–523, 2020.
- [16] Abdolmajid Hayati and Firooz Mohmedi. The Effect of Films With and Without Subtitles on Listening Comprehension of Efl Learners. *British Journal of Educational Technology*, Vol. 42, No. 1, pp. 181–192, 2011.
- [17] Setsuko Mori. If You Build It, They Will Come: From a “Field of Dreams” to a More Realistic View of Extensive Reading in an EFL Context. *Reading in a foreign language*, Vol. 27, pp. 129–135, 2015.
- [18] Eileen Kay Blau. The Effect of Syntax, Speed, and Pauses on Listening Comprehension. *TESOL Quarterly*, Vol. 24, No. 4, pp. 746–753, 1990.
- [19] Katsuya Fujii and Jun Rekimoto. SubMe: An Interactive Subtitle System with English Skill Estimation Using Eye Tracking. In *Proceedings of the 10th Augmented Human International Conference 2019*, pp. 1–9. Association for Computing Machinery, 2019.
- [20] James Milton and Thomai Alexiou. Vocabulary Size and the Common European Framework of Reference for Languages. *Vocabulary Studies in First and Second Language Acquisition*, pp. 194–211, 2009.
- [21] Lars Stenius Staehr. Vocabulary Size and the Skills of Listening, Reading and Writing. *The Language Learning Journal*, Vol. 36, No. 2, pp. 139–152, 2008.
- [22] Stuart Webb and Michael P. H. Rodgers. Vocabulary Demands of Television Programs. *Language Learning*, Vol. 59, pp. 335–366, 2009.
- [23] Michael PH Rodgers. *English Language Learning Through Viewing Television: an Investigation of Comprehension, Incidental Vocabulary Acquisition, Lexical Coverage, Attitudes, and Captions*. PhD thesis, Victoria University of Wellington, 2013.

- [24] Geòrgia Pujadas and Carmen Muñoz. Extensive Viewing of Captioned and Subtitled Tv Series: a Study of L2 Vocabulary Learning by Adolescents. *The Language Learning Journal*, Vol. 47, No. 4, pp. 479–496, 2019.
- [25] Stuart Webb. Using Glossaries to Increase the Lexical Coverage of Television Programs. *Reading in a foreign language*, Vol. 22, pp. 201–221, 2010.
- [26] Stuart Webb. Pre-learning Low-frequency Vocabulary in Second Language Television Programmes. *Language Teaching Research*, Vol. 14, pp. 501–515, 2010.
- [27] Paul Markham. Captioned Videotapes and Second-Language Listening Word Recognition. *Foreign language annals*, Vol. 32, pp. 321–328, 1990.
- [28] Francisca M Ivone and Willy A Renandya. Extensive Listening and Viewing in ELT. *Teflin Journal*, Vol. 30, No. 2, pp. 237–256, 2019.
- [29] Mina Lee, Beverly Roskos, and David R. Ewoldsen. The Impact of Subtitles on Comprehension of Narrative Film. *Media Psychology*, Vol. 16, No. 4, pp. 412–440, 2013.
- [30] Francesca Bianchi and Tiziana Ciabattoni. Captions and Subtitles in Efl Learning: an Investigative Study in a Comprehensive Computer Environment. In *Didactas to Ecolingua: an Ongoing Research Project on Translation and Corpus Linguistics*, pp. 69–90, 2008.
- [31] Stuart Webb and Michael P. H. Rodgers. The Lexical Coverage of Movies. *Applied Linguistics*, Vol. 30, No. 3, pp. 407–427, 2009.
- [32] Stuart Webb. A Corpus Driven Study of the Potential for Vocabulary Learning Through Watching Movies. *International Journal of Corpus Linguistics*, Vol. 15, pp. 497–519, 2010.
- [33] Sathaporn "Hubert" Hu and Wesley J. Willett. Kalgan: Video Player for Casual Language Learning. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–6. Association for Computing Machinery, 2018.
- [34] Nathan Sakunkoo and Pattie Sakunkoo. Gliflix: Using Movie Subtitles For Language Learning. In *Proceedings of the 26th Symposium on User Interface Software and Technology*, 2013.
- [35] Kousuke Sugai, Shigeru Yamane, and Kazuo Kanzaki. The Time Domain Factors Affecting EFL Learners' Listening Comprehension: a Study on Japanese EFL Learners. *ARELE: Annual Review of English Language Education in Japan*, Vol. 27, pp. 97–108, 2016.
- [36] Kuno Kurzhals, Fabian Göbel, Katrin Angerbauer, Michael Sedlmair, and Martin Raubal. *A View on the Viewer: Gaze-Adaptive Captions for Videos*, pp. 1–12. Association for Computing Machinery, 2020.

- [37] Peter D MacIntyre and Robert C Gardner. The Subtle Effects of Language Anxiety on Cognitive Processing in the Second Language. *Language learning*, Vol. 44, No. 2, pp. 283–305, 1994.
- [38] Peter D MacIntyre and Robert C Gardner. Anxiety and Second-language Learning: Toward a Theoretical Clarification. *Language learning*, Vol. 39, No. 2, pp. 251–275, 1989.
- [39] Yasser Teimouri, Julia Goetze, and Luke Plonsky. Second Language Anxiety and Achievement: a Meta-analysis. *Studies in Second Language Acquisition*, Vol. 41, No. 2, pp. 363–387, 2019.
- [40] Peter D MacIntyre. How Does Anxiety Affect Second Language Learning? A Reply To Sparks And Ganschow. *The modern language journal*, Vol. 79, No. 1, pp. 90–99, 1995.
- [41] Tammy S Gregersen. Nonverbal Cues: Clues To The Detection Of Foreign Language Anxiety. *Foreign language annals*, Vol. 38, No. 3, pp. 388–400, 2005.
- [42] 富士通株式会社. パソコンを使う時の姿勢. <https://www.fujitsu.com/jp/about/businesspolicy/tech/design/ud/vdt/index-page3.html>.
- [43] David Bohlke, Paul Dummett, Lewis Lansford, and Helen Stephenson. *Keynote, American English*. CENGAGE Learning, 2016.
- [44] 岡田徹也, 土橋健一郎, 加藤正人, 松園保則. 読むだけでわかる! TOEFL テストガイドブック. アゴス・ジャパン, 2014.
- [45] Tadas Baltrušaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. OpenFace 2.0: Facial Behavior Analysis Toolkit. In *13th IEEE International Conference on Automatic Face Gesture Recognition*, pp. 59–66, 2018.
- [46] Erroll Wood, Tadas Baltrušaitis, Xucong Zhang, Yusuke Sugano, Peter Robinson, and Andreas Bulling. Rendering of Eyes for Eye-Shape Registration and Gaze Estimation. pp. 3756–3764, 2015.
- [47] Tadas Baltrušaitis, Marwa Mahmoud, and Peter Robinson. Cross-dataset Learning and Person-specific Normalisation for Automatic Action Unit Detection. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, Vol. 06, pp. 1–6, 2015.
- [48] Paul Ekman. Facial Expression and Emotion. *American psychologist*, Vol. 48, No. 4, pp. 384–392, 1993.
- [49] Jeffrey Cohn, Zara Ambadar, and Paul Ekman. *Observer-Based Measurement of Facial Expression with the Facial Action Coding System*, pp. 2–53. Oxford University Press Series in Affective Science, 2007.

- [50] Mehdi Ghayoumi and Arvind Bansal. Unifying Geometric Features and Facial Action Units for Improved Performance of Facial Expression Analysis. 2016.
- [51] Y. I. Tian, T. Kanade, and J.F. Cohn. Recognizing Action Units for Facial Expression Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pp. 97–115, 2001.
- [52] Paul Ekman and Wallace V Friesen. *Facial Action Coding System*. Palo Alto, CA: Consulting Psychologists Press, 1978.
- [53] Jacob Cohen. Statistical Power Analysis for the Behavioral Sciences. *Academic Press*, 2013.
- [54] 小屋多恵子. 初級英語学習者の聴解に与える発話速度調整の効果. 小金井論集, Vol. 13, pp. 11–30, 2017.
- [55] Xinlei Zhang, Takashi Miyaki, and Jun Rekimoto. WithYou: Automated Adaptive Speech Tutoring With Context-Dependent Speech Recognition. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. Association for Computing Machinery, 2020.
- [56] 高倉祐樹, 大槻美佳, 中川賀嗣, 大澤朋史, 谷川緑野. 言語性短期記憶のメカニズムとその障害について. 高次脳機能研究, Vol. 31, No. 4, pp. 411–421, 2011.
- [57] Jon Gillick. Laughter-Detection, 2021. <https://github.com/jrgillick/laughter-detection> (最終閲覧日: 2022年1月21日).
- [58] Octavio Arriaga, Matias Valdenegro-Toro, Mohandass Muthuraja, Sushma Devaramani, and Frank Kirchner. Perception for Autonomous Systems, 2020.
- [59] Octavio Arriaga, Matias Valdenegro-Toro, and Paul Plöger. Real-time Convolutional Neural Networks for Emotion and Gender Classification, 2017.
- [60] 西田直人, 野崎陽奈子, 志築文太郎. 表情に基づく再生速度自動調節機能を備えた外国語学習支援の提案. 第29回インタラクティブシステムとソフトウェアに関するワークショップ, 2021.
- [61] Software Freedom Conservancy. Selenium, 2021. <https://www.selenium.dev/ja/> (最終閲覧日: 2022年1月21日).
- [62] Council of Europe. *Common European framework of reference for languages: Learning, teaching, assessment*. U.K: Press Syndicate of the University of Cambridge, 2001.
- [63] Geòrgia Pujadas and Carmen Muñoz. Examining Adolescent EFL Learners’ TV Viewing Comprehension Through Captions And Subtitles. *Studies in Second Language Acquisition*, Vol. 42, No. 3, p. 551–575, 2020.

- [64] Veronica Benigno and John HAL De Jong. *Linking vocabulary to the CEFR and the Global Scale of English: A psychometric model*, pp. 8–29. EALTA-European Association for Language Testing and Assessment, University of Jyväskylä, Centre for Applied Language Studies, 2019.
- [65] R.J. Tannenbaum and E. Wylie. *Linking English-Language Test Scores Onto the Common European Framework of Reference: An Application of Standard-Setting Methodology*. Technical report, ETS, 2008.
- [66] English-Corpora.org. *Corpus of Contemporary American English*, 2021. <https://www.english-corpora.org/coca/> (最終閲覧日：2022年1月21日) .
- [67] Gwan-Hyeok Im and Liying Cheng. *The Test of English for International Communication (TOEIC®)*. *Language Testing*, Vol. 36, No. 2, pp. 315–324, 2019.
- [68] G Frederic Kuder and Marion W Richardson. *The Theory of The Estimation of Test Reliability*. *Psychometrika*, Vol. 2, No. 3, pp. 151–160, 1937.
- [69] Sandra G. Hart and Lowell E. Staveland. *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*. In Peter A. Hancock and Najmedin Meshkati, editors, *Human Mental Workload*, Vol. 52, pp. 139–183. North-Holland, 1988.
- [70] John Brooke. *SUS: A “Quick and Dirty” Usability Scale*, pp. 189–194. London: Taylor and Francis, 1996.
- [71] 佐藤和人, 大津宏亮, 間所洋和, 門脇さくら. 意図的な表情表出に及ぼす心理的ストレス要因の分析. *情報科学技術フォーラム講演論文集*, 第 12 巻, pp. 21–28, 2013.
- [72] Aaron Bangor, Philip Kortum, and James Miller. *Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale*. *Journal of Usability Studies*, Vol. 4, No. 3, pp. 114–123, 2009.
- [73] Roy C Major, Susan M Fitzmaurice, Ferenc Bunta, and Chandrika Balasubramanian. *Testing the Effects of Regional, Ethnic, and International Dialects of English on Listening Comprehension*. *Language learning*, Vol. 55, No. 1, pp. 37–69, 2005.
- [74] イヴァンボイツォフ, 白山利信, 加藤百合. 日本人ロシア語学習者の言語的誤りとその矯正用の教材開発に関する総合的研究, 2021–2025. <https://kaken.nii.ac.jp/ja/grant/KAKENHI-PROJECT-21K00647/>.
- [75] 白山利信. 平成 29 年度～30 年度 (2017～2018 年度) CEGLOC 外国語教育部門活動報告 . 2019.

## 著者論文リスト

著者は学類生期間にて学会発表を行った。下記に発表の一覧を記す。

1. 西田直人, 野崎陽奈子, 志築文太郎. 表情に基づく再生速度自動調節機能を備えた外国語学習支援の提案. 第 29 回インタラクティブシステムとソフトウェアに関するワークショップ, 2021. 2 pages.
2. Yuga Tsukuda, Naoto Nishida, Jun Lu, and Yoichi Ochiai. Insect-Computer Hybrid Speaker: Speaker using Chirp of the Cicada Controlled by Electrical Muscle Stimulation. The ACM CHI 2021 Workshop on Design and Creation of Inclusive User Interactions Through Immersive Media. 2021. 5pages.

## 付録A 予備実験にて用いた各種書類

本付録にて、予備実験に関する書類としてアンケートを以下に示す。

### A.1 実験後アンケート

### 予備実験

 nishida@iplab.cs.tsukuba.ac.jp (共有なし)   
[アカウントを切り替える](#)

\*必須

名前 \*

回答を入力

英語はあなたの母語ですか \*

はい

いいえ

外国語の動画を視聴している、または外国語で行われる会議や授業に参加している時、理解が追いつかないと感じたことはありますか (例: 学部の英語の授業、洋画の視聴、国際会議のプレゼン動画の視聴、国際会議への参加時) \*

ある

ない

母国語の動画を視聴している、または会議や授業に参加している時、理解が追いつかないと感じたことはありますか (例: 難しい数学の授業)

ある

ない

分からない(自分の理解が追いつかない)時取る癖はありますか。ある場合は軽く記述してください(貧乏ゆすり, 視線が逸れる, 表情が変化するなど)\*

回答を入力 \_\_\_\_\_

外国語のコンテンツを視聴している際に困ることはなんですか

外国語を話す速度が速すぎる

話者の訛りがひどい

語彙が分からない

その他: \_\_\_\_\_

外国語能力の指標を教えてください(例: TOEFL, TOIEC, 英検などのスコア, CEFRスコア, もしくは体感でどの動画が何%ほど理解できるか)

回答を入力 \_\_\_\_\_

日常生活で英語はどれくらい使いますか(時間/週)\*

~2h

2h~5h

5h~10h

10h~15h

15h~20h

20h~25h

25h~40h

日常生活で英語はどれくらい使いますか(時間/週) \*

- ~2h
- 2h~5h
- 5h~10h
- 10h~15h
- 15h~20h
- 20h~25h
- 25h~40h
- 40h~60h
- 60h~80h
- 80h~

ビデオの中において理解が難しい場面はありましたか（印象に残る場面だけでも） \*

回答を入力

教材の動画はどれくらい理解できましたか \*

- 1    2    3    4    5    6    7
- 全くわからない                        全部わかった

問題文が表示されている時間の長さはどうでしたか \*

問題文が表示されている時間の長さはどうでしたか\*

1 2 3 4 5 6 7  
短い ○ ○ ○ ○ ○ ○ ○ 長い

分からない時にする表情は次のうちどれですか\*

- 怒り
- 嫌悪
- 恐れ
- 笑顔
- 悲しみ
- 驚き
- 無表情（意図的に）
- 無表情（無意識的に）
- その他: \_\_\_\_\_

その他気になること、気づいたことはありましたか\*

回答を入力 \_\_\_\_\_

送信

フォームをクリア

## 付録B 本実験にて用いた各種書類

本付録にて，実験に関する書類として実験前アンケート，実験中アンケート，および実験後アンケートを以下に示す．

## B.1 実験前アンケート

### 本実験①→②

 nishida@iplab.cs.tsukuba.ac.jp (共有なし)   
[アカウントを切り替える](#)

**\*必須**

名前 \*

回答を入力

メールアドレス \*

回答を入力

性別 \*

男

女

その他: \_\_\_\_\_

年齢 (20歳なら「20」) \*

回答を入力

英語はあなたの母語ですか \*

はい

いいえ

日常生活で英語はどれくらい使いますか(時間/週) \*

~2h

2h~5h

5h~10h

10h~15h

15h~20h

20h~25h

25h~40h

40h~60h

60h~80h

80h~

外国語能力の指標を教えてください(例：TOEFL, TOIEC, 英検などのスコア, CEFRスコアなど) \*

回答を入力

グループ \*

A

c

フォームをクリア

## B.2 実験中アンケート：語彙テスト 1

動画視聴前にもこのアンケートと同様の語彙クイズに答える。ただし、動画視聴前はNASA-TLXの項目はない。

本実験①事前テスト

単語の意味を教えてください

toe

- くるぶし
- 中指
- つま先
- 手の甲

punchline

- 顔面
- 弱点
- 笑いどころ
- ラインエ

carrot

- にんじん
- ビーガン
- ロケット
- カカロット

attach

- 話す
- 投げる
- 落とす
- くっつける

attach

- 話す
- 投げる
- 落とす
- くっつける

sew

- 落とす
- 歩く
- 縫う
- 曲げる

jitter

- イライラ
- 信号の揺らぎ
- バター
- ゴミ

applause

- 塩酸
- アプリ
- 賞賛
- 薔薇

attitude

attitude

- 経度
- 緯度
- 温度
- 態度

appearance

- 外見
- 出現
- 嗜好
- ペンギン

pound

- 顔面
- 重さの単位
- パウンドケーキ
- 沼

gassy

- おしゃべりな
- 気球の
- 膨大な
- 破裂音の

take something apart

**take something apart**

- ~をアパートに住ませる
- ~をなじる
- ~を離す, 分ける
- ~をなぐる

**steal**

- スチール (金属)
- 盗む
- 羨む
- 蔑む

**separate**

- セリーグの
- 分ける
- さわぐ
- 舞台の

**insane**

- 正しい
- 面白い
- 正気でない
- 背が高い

**hanging body**

ぶた

目立つ人間

首吊り死体

separate

分ける

セリーグの

さわぐ

舞台の

identical twin

二卵性双生児

厳密な分身

一卵性双生児

そっくりさん

NASA-TLX : <https://www.keithv.com/software/nasatlx/nasatlx-ja.html>  
<https://www.keithv.com/software/nasatlx/nasatlx-ja.html>

NASATLXの結果を貼り付けてください

回答を入力

### B.3 実験中アンケート：語彙テスト2

動画視聴前にもこのアンケートと同様の語彙クイズに答える。ただし、動画視聴前はNASA-TLXの項目はない。

本実験②事前テスト

単語の意味を教えてください

generous

- 変な
- ラッキーな
- 寛大な
- 暑い

weird

- 散らばっている
- 気持ち悪い
- 美しい
- 大変

fool around

- ワクワクさせる
- 拡大する
- ポコポコにする
- ～と遊ぶ、いちゃつく

go out with somebody

- ～とデートに出かける
- ～と殴り合う
- ～と話し合う

go out with somebody

- ~とデートに出かける
- ~と殴り合う
- ~と話し合う
- ~と蹴り合う

storage room

- ロッカー
- 倉庫
- データ管理場所
- リビング

trifle

- トライフル(デザートの名前)
- ライフル
- 変なもの
- メインディッシュ

playboy

- 野球少年
- プレイボーイ (雑誌の名前)
- ヤンキー
- 不審者

porch swing

porch swing

- ホームラン
- 密猟
- 翼
- ベンチ

girly

- 女々しい
- 曲がっている
- ギリギリと音がする
- ラッキーな

masculine

- 大虐殺の
- 男らしい
- 臭い
- マスカラの

terrific

- 仰々しい
- 大陸の
- マグマの
- ものすごい

get fired

get fired

- 発砲する
- 放火する
- クビになる
- 火だるまになる

cheat on somebody

- ~を浮気する
- ~をカンニングする
- ~をいじる
- ~を笑う

divorce

- 反抗
- 反発
- 結婚
- 離婚

messy

- めちゃくちゃな
- 重い
- 巧妙な
- ラッキーな

imply

- 大陸の
- マグマの
- ものすごい

imply

- ほのめかす
- 殺す
- 助ける
- しゃべる

divorce

- 反抗
- 反発
- 結婚
- 離婚

NASA-TLX : <https://www.keithv.com/software/nasatlx/nasatlx-ja.html>  
<https://www.keithv.com/software/nasatlx/nasatlx-ja.html>

NASATLXの結果を張ってください

回答を入力

戻る

次へ

フォームをクリア

## B.4 実験中アンケート：内容把握に関するクイズ1

クイズ① name:

What did the woman with vodka and cranberry juice do?

- a) she didn't do anything
- b) she took a photo of her husband
- c) she killed her husband
- d) she shot her husband

What is a true sentence?

- a) The doctor needed to sew the vegetable as a toe
- b) A woman brought a toe
- c) A woman brought a tarot
- d) A doctor was asked to sew a toe

What is the punchline?

- a) that she said camera adds 10 pounds
- b) that he guessed she was too skinny
- c) that he implied she was fatter than she said she was
- d) that he implied she had 10 pounds fatter before

What did the man want the woman to do?

- a) hold a fashion show
- b) serve him some coffee
- c) get the hangers together
- d) take the hangers apart

What did he want?

- a) he wanted an applause
- b) he wanted a recipe of jitter
- c) he wanted a recipe of tiramisu
- d) he wanted a grandma's attention

What did he argue?

- a) that they didn't know her name
- b) that they had a mistake on her appearance
- c) that they needed some extra information
- d) that they had a mistake on her attitudes

What is a true sentence?

- a) he is eating the woman's food
- b) she is living in this room
- c) he is living in this room
- d) he ate the Ross's chips

Why does the man need to be cool?

- a) because the grandma doesn't like him
- b) because the grandma doesn't know him
- c) because the grandma probably know about him
- d) because the grandpa doesn't support him

What is a true sentence?

- a) Carl is Joey's identical twin
- b) Carl is nothing to do with Joey
- c) Someone employed Carl and Joey
- d) Carl is none of the relatives of Joey

## B.5 実験中アンケート：内容把握に関するクイズ2

クイズ② name:

What is the punchline?

- a) he said she is generous
- b) he said sweet words
- c) he said she is strange
- d) she didn't like the word he said

Who is the woman they mentioned?

- a) a woman who went out with the man
- b) a woman who the man can recall
- c) a woman who the man played with
- d) a woman who loved him

What is the punchline?

- a) he get his foot a pedicure
- b) she implied his attitude seemed girly
- c) he showed a masculine gesture
- d) he is not a weird

What is a true sentence?

- a) Ross got divorced in LA
- b) Ross stole the playboy
- c) Monica and Chandler are not living together
- d) she did easily understand the information

What is a true sentence?

- a) The man's wife cheated on him
- b) The man's wife is getting divorced
- c) The man's ex-girlfriend is getting married
- d) The man's best friend is getting married

What does the man compliment about?

- a) The priest said the right name
- b) The wedding was messy
- c) The wedding was not the first wedding
- d) The wedding was full of characters

What does the man imply?

- a) Emma is a good woman
- b) Emma is a excessively caring hugsy
- c) Hugsy loves Emma
- d) Hugsy is a wonderful kid

What will the man do?

- a) Pretend to be a nice guy
- b) Behave as usual to a modest level
- c) be shot down with the woman's rifle
- d) accuse her of the disorder

What does the woman suggest?

- a) he should play the keyboard again
- b) he should hold a private lesson
- c) he should play in public
- d) he should play in a pub

## B.6 実験中アンケート：SUS

System Usability Scale

本システムに関して、以下の10項目を評価してください。

本システムを頻繁に使いたいと思う。

	1	2	3	4	5	
そう思わない	<input type="radio"/>	そう思う				

本システムは不必要に複雑だと思う

	1	2	3	4	5	
そう思わない	<input type="radio"/>	そう思う				

本システムは使いやすいと思う

	1	2	3	4	5	
そう思わない	<input type="radio"/>	そう思う				

本システムを使用するには、詳しい人のサポートが必要だと思う

	1	2	3	4	5	
そう思わない	<input type="radio"/>	そう思う				

本システムは様々な機能が上手く統一されていると思う

	1	2	3	4	5	
そう思わない	<input type="radio"/>	そう思う				

本システムには一貫性がないと思う

1 2 3 4 5  
そう思わない ○ ○ ○ ○ ○ そう思う

多くの人が本システムの使い方をすぐに習得できると思う

1 2 3 4 5  
そう思わない ○ ○ ○ ○ ○ そう思う

本システムはとても操作しづらいと思う

1 2 3 4 5  
そう思わない ○ ○ ○ ○ ○ そう思う

本システムを使いこなせる自信がある

1 2 3 4 5  
そう思わない ○ ○ ○ ○ ○ そう思う

本システムを使い始める前に、多くのことを学ぶ必要がある

1 2 3 4 5  
そう思わない ○ ○ ○ ○ ○ そう思う

戻る

次へ

フォームをクリア

## B.7 実験後アンケート

本実験 (2 / 2)

デバイスを用いた時に動画コンテンツを楽しむ事はできましたか \*

1 2 3 4 5 6 7

楽しむことができなかった         楽しむことができました

デバイスを用いなかった時に動画コンテンツを楽しむ事はできましたか \*

1 2 3 4 5 6 7

楽しむことができなかった         楽しむことができました

デバイスを用いた時の動画の速度調整は適切でしたか \*

1 2 3 4 5 6 7

適切でなかった         適切だった

表情検出から自動で再生速度が変わるシステムを使用した際、自分で速度を調整した方がいいと感じた場面がありましたか \*

1 2 3 4 5 6 7

なかった         あった

上記の質問で「5, 6, 7」を選んだ場合(自分で速度を調整した方がいいと感じた場合)理由を書いてください

回答を入力

今回の動画コンテンツ (friendsというドラマ) は以前に試聴したことがありますか \*

上記の質問で「5, 6, 7」を選んだ場合(自分で速度を調整した方がいいと感じた場合)理由を書いてください

回答を入力

今回の動画コンテンツ (friendsというドラマ) は以前に試聴したことがありますか \*

- はい
- いいえ

よりよく理解できた動画は①と②のどちらですか \*

- ①
- ②

動画内で最後までわからなかった単語があれば書いてください

回答を入力

その他気になること、気づいたことはありましたか \*

回答を入力

戻る

送信

フォームをクリア