

クラウドソーシングを用いた
リアルタイム手話文字通訳システムに
関する研究

令和 3 年度

筑波技術大学大学院技術科学研究科

産業技術学専攻

田中 康平

目次

第1章	序論	1
1.1.	研究背景	1
1.2.	研究目的	6
1.3.	本論文の構成	6
第2章	関連研究	9
2.1.	音声文字起こしシステム	9
2.2.	SCRIBE	9
2.3.	クラウドソーシングを用いた手話文字通訳	10
2.4.	本研究で実装するシステム	12
第3章	手話文字通訳システム	13
3.1.	システムの概要	13
3.2.	システムアーキテクチャ	13
3.3.	管理画面	14
3.4.	通訳画面	15
3.4.1.	通訳画面の動画と入力枠	15
3.4.2.	動画操作機能と早送り	15
3.4.3.	指示枠	16
3.4.4.	前後入力文	17
3.4.5.	動画サイズ, 文字の大きさ, ワーカの人数	17
3.5.	システムの不具合対処方法	17
3.6.	字幕画面	18
3.7.	システム全体の処理	18
第4章	手話文字通訳の予備実験	21
4.1.	実験目的	21
4.2.	実験環境と条件	21
4.3.	実験結果	22
4.3.1.	各ワーカのタイピング速度と通訳時間	22
4.3.2.	通訳区間の間の文の欠落と重複のある通訳区間の割合	23
4.3.3.	各ワーカが担当した通訳区間	24
4.3.4.	主観評価結果	25

4.4.	考察	25
4.4.1.	タイピング速度と通訳時間	25
4.4.2.	連係入力	26
4.4.3.	遅延を取り戻すための早送り	26
4.4.4.	通訳の楽しさと必要人数	26
4.5.	まとめ	27
第5章	手話文字通訳の本実験	29
5.1.	改良した通訳画面	29
5.2.	実験目的	30
5.3.	実験方法	30
5.4.	実験参加者	30
5.5.	手話動画と通訳区間の長さ	31
5.6.	システムユーザビリティスケール(SUS) [22]	33
5.7.	アンケート	33
5.7.1.	手話文字通訳システムの慣れ	33
5.7.2.	前後入力文を確認したか	33
5.7.3.	再生倍率 1.5 倍速で遅延を取り戻す方法	33
5.7.4.	通訳の楽しさ	34
5.8.	通訳品質の評価方法	34
5.9.	実験結果	35
5.9.1.	手話文字通訳システムの慣れ	35
5.9.2.	通訳時間と必要人数	39
5.9.3.	通訳回数	40
5.9.4.	通訳品質と通訳区間の間の文の欠落と重複のある通訳区間の割合	41
5.9.5.	前後の入力文を確認したかの主観評価結果	41
5.9.6.	再生倍率 1.5 倍速は速く感じたかの主観評価結果	42
5.9.7.	動画操作機能の主観評価結果	43
5.9.8.	通訳の楽しさの主観評価結果	44
5.9.9.	SUS による評価	45
5.10.	考察	46
5.10.1.	手話文字通訳の慣れ	46
5.10.2.	連係入力の慣れ	47
5.10.3.	動画操作の慣れ	47
5.10.4.	通訳時間	48
5.10.5.	必要人数	49
5.10.6.	通訳品質	49

5. 10. 7. 関係入力	50
5. 10. 8. 通訳区間の間の文の欠落と重複の発生理由	51
5. 10. 9. 再生倍率 1.5 倍速で遅延を取り戻す方法の主観評価結果	51
5. 10. 10. 通訳回数と通訳の楽しさ	52
5. 10. 11. SUS を用いた主観評価結果	53
5. 10. 12. まとめ	53
第 6 章 結論	55
6. 1. まとめ	55
6. 2. 今後の課題	56
謝辞	59
参考文献	61
本研究に関する成果・発表等	65
付録	67
A1 開発環境	67
A2 第5章の本実験のアンケート	67
A2.1 手話文字通訳システムの慣れについてのアンケート	67
A2.2 手話文字通訳実験についてのアンケート	69

図目次

図 1.1 手話通訳者.....	3
図 1.2 文字通訳者.....	3
図 1.3 2段階変換.....	4
図 1.4 クラウドソーシングを用いた手話文字通訳システム.....	4
図 1.5 通訳区間の一定時間区切りと通訳区間の間の重複.....	4
図 1.6 ワークインタフェース.....	5
図 1.7 通訳区間の間の文の欠落と重複.....	5
図 1.8 入力完了後に遅延が発生する.....	5
図 1.9 遅延を取り戻すために早送りでライブの再生位置まで追いつく.....	6
図 1.10 前後のワークの入力途中文を確認する連係入力.....	6
図 2.1 SCRIBE.....	10
図 2.2 クラウドソーシングを用いた手話文字通訳システム.....	11
図 2.3 グループの入力結果の評価方法.....	12
図 3.1 手話文字通訳システムのアーキテクチャ.....	14
図 3.2 管理画面.....	15
図 3.3 通訳画面.....	17
図 3.4 字幕画面.....	18
図 3.5 手話文字通訳システムの処理の流れ.....	20
図 4.1 通訳時間とタイピング速度の散布図.....	23
図 4.2 通訳区間の間の文の欠落や重複のある通訳区間の割合.....	24
図 4.3 ワークが通訳した通訳区間.....	24
図 5.1 通訳画面.....	29
図 5.2 手話文字通訳の慣れの平均と標準偏差.....	37
図 5.3 連係入力の慣れの平均と標準偏差.....	38
図 5.4 動画操作機能の慣れの平均と標準偏差.....	38
図 5.5 各ワークの平均通訳時間のヒストグラム.....	40
図 5.6 「この中で使った機能はどれですか？」の主観評価結果.....	44

表目次

表 4.1 タイピング速度と通訳時間の平均と標準偏差	22
表 4.2 各ワーカの主観評価結果	25
表 5.1 各実験の実験参加者数とタイピング速度の平均と標準偏差	31
表 5.2 各講演者の手話の速さ	32
表 5.3 通訳時間とワーカ人数と必要人数	39
表 5.4 1回目から5回目の通訳回数の平均と標準偏差	40
表 5.5 通訳品質と通訳区間の間の文の欠落率と重複率	41
表 5.6 前後入力文を確認したかの主観評価結果	42
表 5.7 再生速度 1.5 は速く感じたかの平均と標準偏差	43
表 5.8 通訳の楽しさの主観評価結果	45
表 5.9 SUS の評価結果	45
表 5.10 SUS の質問項目ごとの平均点と標準偏差	46

筑波技術大学

修士（工学）学位論文

第1章 序論

1. 1. 研究背景

ろう者・難聴者は一般的に音声の聞き取りが困難であり、主なコミュニケーション手段として手話言語や文字言語を用いている。しかし、聴者による講演では、音声で発話することが一般的である。そのため、音声の聞き取りが困難な者は聴者と同等の情報量を得ることは難しい。通訳者が音声を手話または文字に通訳することで、ろう者・難聴者でも情報を得ることができる。ただし、通訳者が正確に通訳できない場合、講演内容の理解が困難となる。また、通訳の遅延が大きいほど講演の進行状況の把握が困難となる。したがって、正確にかつ速く通訳するために相当量の訓練を行った手話通訳者や文字通訳者が情報を保障する(以下、情報保障)ことが一般的である。聴者による講演での現状の情報保障として2つの方法がある。1つ目は、図1.1のように、講演者の発話を手話通訳者が聞き取りながら手話に通訳し、手話の分かるろう者・難聴者が手話を読み取る方法がある。2つ目は、図1.2のように、講演者の発話を文字通訳者が聞きながらPCを用いて文字に通訳し、字幕に表出してろう者・難聴者に見てもらう方法がある。手話通訳は手話を読める者、文字通訳は日本語が分かる者に対して有効である。

手話話者による講演では、手話を読み取れない聴者のために、手話通訳者が手話を音声に通訳する方法が一般的である。ところが、手話を読めないろう者・難聴者は音声の聞き取りが困難であるため、文字への通訳が求められる。そこで、手話講演による現状の情報保障として、図1.3のように、講演者の手話を手話通訳者が読み取りながら音声に通訳し、文字通訳者が音声を聞きながら文字通訳する方法がある。文字通訳者は一般的に手話言語を習得していないため、手話を直接文字に通訳する技術を持たない者が多い。手話通訳者が講演者の手話を音声に変換することで、文字通訳者は音声を文字に通訳できるようになる。しかし、2段階通訳のため、各通訳者には別々の料金を支払うことになり、聴者による講演の情報保障と比べて通訳派遣に支払う金額が高い。また、文字は最後に通訳されるため、字幕表示の遅延が著しい。この課題を解決するために、本研究では、図1.4のように、クラウドソーシングを用いて複数の非専門家ながら手話を読めるボランティアのワーカが手話を直接文字に通訳する方法(以下、手話文字通訳)を検討した。これまでに、ろう者・難聴者は支援される側であったが、この方法であれば、手話の読み取れるろう者・難聴者が手話を読めない者を支援するための環境を提供することができる。

手話文字通訳の場合、手話を日本語に通訳する必要があるため、通訳しながら文字入力する技能を要する。また、音声は聴覚言語のため、聴きながら文字起こしが可能であるが、手話言語は視覚言語であるため、手話と文字を交互に見ながら入力する技能を要する。しかし、ワーカは一般的に通訳者としての特別な訓練を受けていない。経験の少ないワーカは手話と文字を交互に見ながら入力すると手話の見落としが発生しやすい。また、タイピングが遅いワーカは入力が発話に追いつかない時に負担が非常に大きく、字幕表示までの遅延が大きくなる。受講者は、字幕の誤りが多く、字幕表示までの遅延が大きいと講演の進行状況の把握や内容理解が困難となる。ボランティ

アは無償のため、長時間の訓練や非常に負担の大きい作業があると参加意欲が低下する。参加したとしても負担が大きくて楽しめなかった場合、継続して参加しなくなることも想定できる。そこで、短時間の訓練で、楽しめるような手話文字通訳システムを構築できれば、ボランティアの参加意欲を高めることが期待できる。さらに、字幕の誤りや表示の遅れを減らすための、手話文字通訳を行うタスク設計が可能であれば、手話を読み取れない受講者も参加しやすくなる。

クラウドソーシングを介した手話文字通訳に関する研究[10]がある。この研究では、講演者のいる現場にワーカ全員が集まり、ワーカは講演者の手話を読み取りながら役割にあった作業を行う。次に、通訳区間を作成する作業を担当する1名のワーカ(以下、カウンタ係)が意味単位で通訳区間を作成し、1つの通訳区間は1グループ内の複数のワーカ(以下、通訳ワーカ)が担当する。さらに、グループ内の各通訳ワーカは、通訳区間を割り当てられた時、手話文字通訳を行い、複数の入力結果の文章が生成される。最後に、投票を担当するワーカ(以下、投票ワーカ)が1グループの入力結果から通訳品質の高い最適文を選び、字幕に表示する。しかし、通訳区間を意味単位で作成する方法では、通訳区間の時間が一定ではなく、通訳ワーカの作業量や通訳品質はカウンタ係の能力に依存する。また、カウンタ係は常に手話を見なければならないため、負担が大きい。通訳品質を保障するために1グループ3名以上を5グループ以上で行うが、投票ワーカとカウンタ係を含めて人的コストが高い。

我々はこれまでに、非専門家でもできる手話文字通訳のタスクの設計を検討した[1]。カウンタ係をなくすために、図 1.5 のように、手話動画を自動で一定時間の区間(以下、通訳区間)を短く区切り、複数のワーカで分担する。一定時間で区切ると、通訳区間の終わりと次の通訳区間の初めの間(以下、通訳区間の間)に手話表現が途切れることがあり、通訳区間の間の文の欠落が発生する。欠落を防ぐために、通訳区間の間を重複させた。人的コストを低くするために、1つの通訳区間をグループではなく1名で担当する。ワーカインタフェースには図 1.6 のように、手話と文字との視線移動を減らすために事前に保存した手話動画を表示させ、その下に文字を入力する枠を表示した。字幕の誤りを減らすために、投票の方法ではなく、動画操作機能を設けた。手話の見落としを防ぐために巻き戻しができるようにした。発話に追いつかない場合、再生速度を低くしたり、停止したりできるようにした。しかし、このシステムはリアルタイムの情報保障を実現するために必要であるライブ動画配信に対応していないことや、図 1.7 のように、通訳区間の間を重複しても、文の欠落や重複は依然として残っていた。

リアルタイムの情報保障を実現するために、ライブ動画配信を検討する。しかし、ライブ動画の手話文字通訳を行う際に、手話の見落としや発話速度に追いつくための動画操作機能を用いると、図 1.8 のように、入力完了後にライブまでの遅延が大きく発生する。つまり、入力完了後にライブの再生位置から見始めると、それまでの手話発話を確認することができず、文脈の把握が困難となる。これを解決するために、図 1.9 のように、入力完了後、再生速度を上げて早送りでライブの再生位置まで追いつきながら手話を読み取る方法を検討する。通訳品質を保障するには、手話の見落としのみならず、通訳区間の間の文の欠落や重複を防ぐ必要がある。文の欠落や重複が残ると、受

講者は字幕を読みにくいと感じるからである。文の欠落や重複をなくすために、図 1.10 のように、前後のワーカの入力途中文を見ることができるようになって、連係入力を検討する。



図 1.1 手話通訳者



図 1.2 文字通訳者

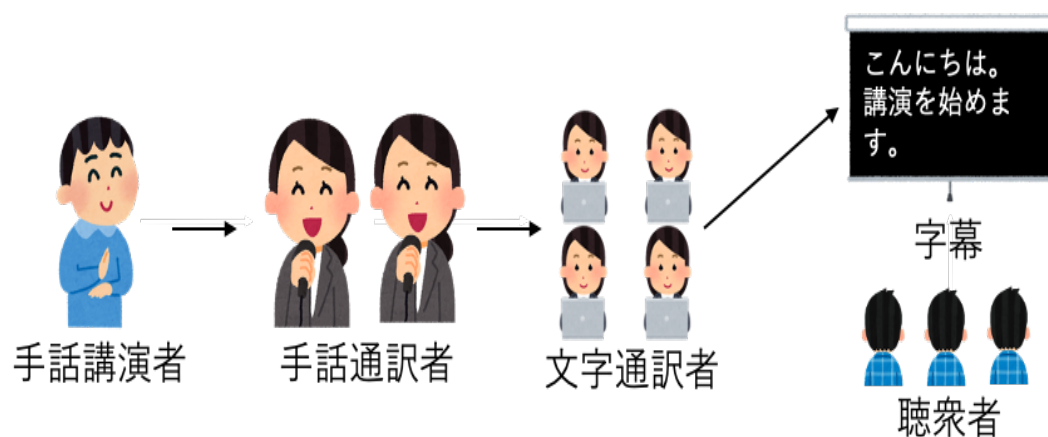


図 1.3 2段階変換



図 1.4 クラウドソーシングを用いた手話文字通訳システム

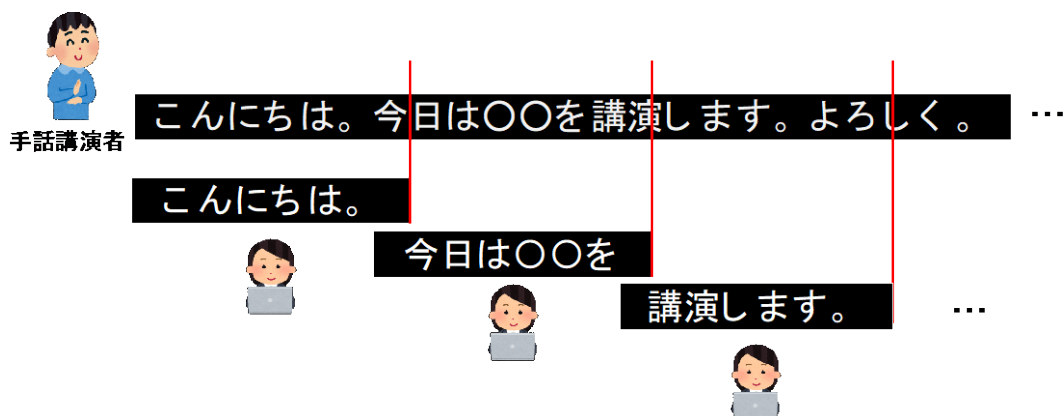


図 1.5 通訳区間の一定時間区切りと通訳区間の間の重複

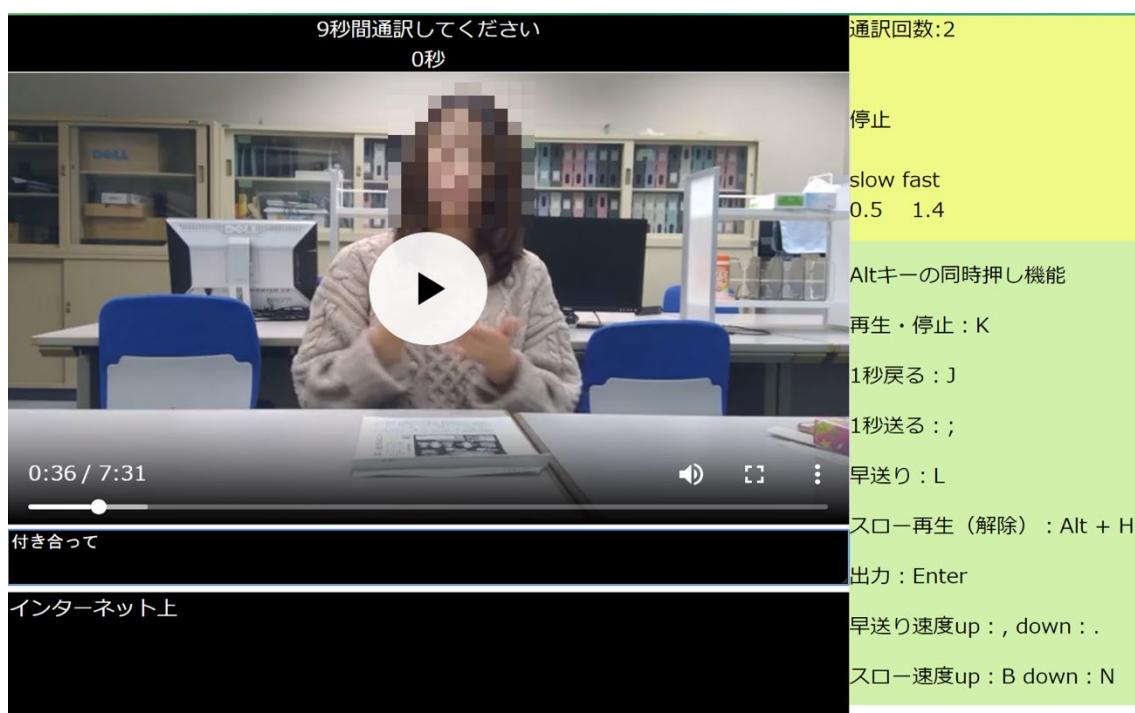


図 1.6 ワークインターフェース

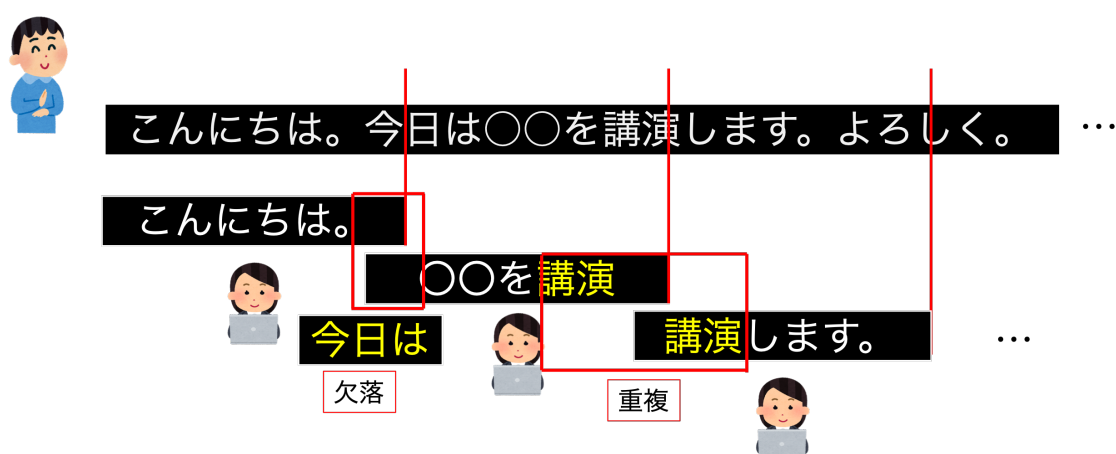


図 1.7 通訳区間の間の文の欠落と重複

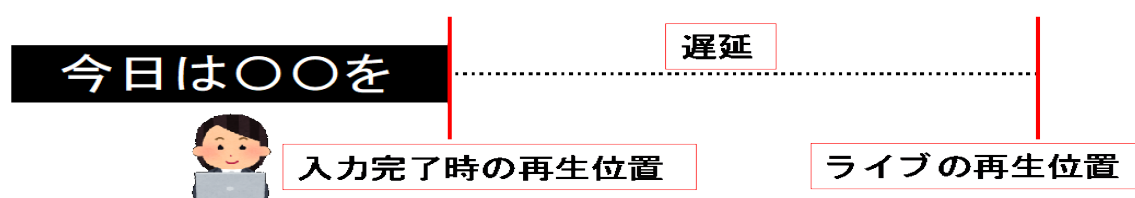


図 1.8 入力完了後に遅延が発生する

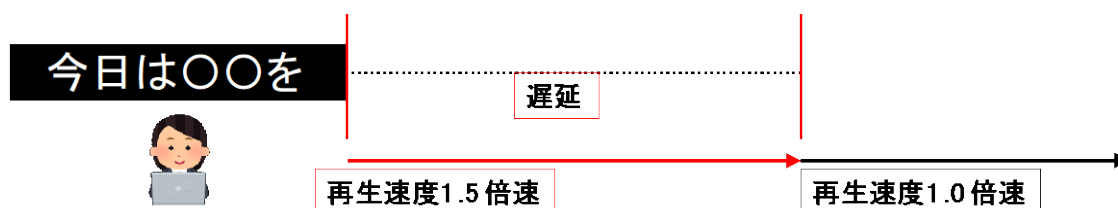


図 1.9 遅延を取り戻すために早送りでライブの再生位置まで追いつく

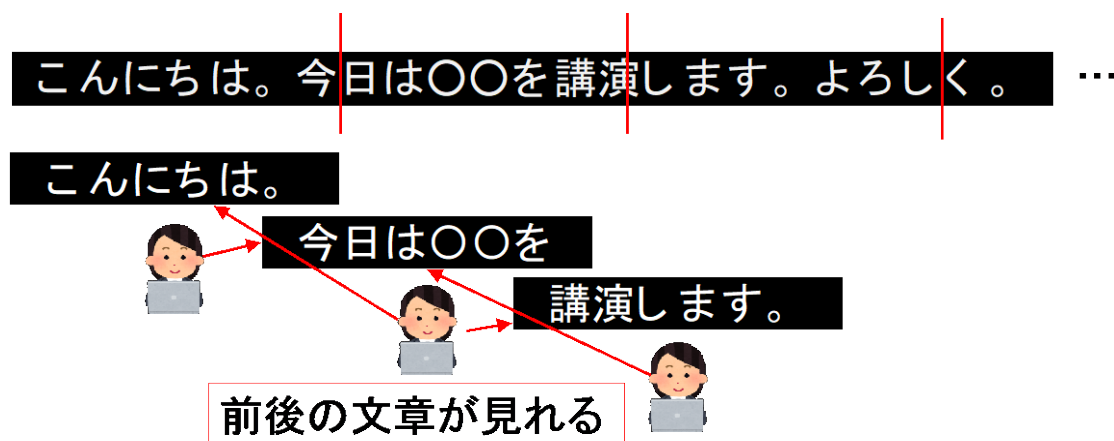


図 1.10 前後のワークの入力途中文を確認する連係入力

1. 2. 研究目的

本研究の目的はクラウドソーシングを用いた複数のボランティアのワークが手話文字通訳を実現するために、提案するシステムの有効性を評価することである。そのために検討する事柄は4つある。1 つ目は、提案するシステムは字幕表示の遅れや誤り、人的コストを減らせるか評価を行う。具体的には、発話開始から字幕表示までの時間(以下、通訳時間)の客観評価や、通訳結果の意味伝達率(以下、通訳品質)の主観評価、必要人数の推定を行う。2 つ目は、字幕を読む受講者のために、通訳区間の間の文の欠落や重複をなくす連係入力を検討し、文の欠落や重複を減らせたか客観評価を行う。3 つ目は、入力完了後、遅延を取り戻すための早送りは早く感じるか主観評価を行う。4 つ目は、経験の少ないボランティアの参加意欲を高めるために、短時間の訓練で参加できつつ、手話文字通訳を楽しめることを評価する。そのために、システムの慣れや通訳の楽しさ、システムユーザビリティスケール(以下、SUS)[22]の主観評価を行う。

1. 3. 本論文の構成

本章では、本研究の研究背景と目的について述べた。第2章では関連研究とその問題点について述べ、本研究の位置付けを明らかにする。第3章では手話文字通訳システムについて述べる。第4章ではプロトタイプシステムを用いた予備実験結果について述べ、考察を行う。第5章では第4

章の実験結果から改良したプロトタイプシステムを用いた本実験結果について考察を述べる。第6章ではまとめと今後の課題について述べる。

第2章 関連研究

2.1. 音声文字起こしシステム

音声を実タイムで文字起こすために、速記者が特殊なキーボードを用いた CART[2]や、一般的な PC を用いた C-Print[3]などが開発されている。日本では、文字通訳者同士で協力して字幕を作成する IPtalk[4]や Web ベースで文字通訳が可能な captiOnline[5]が利用されている。これらのシステムは、長年の訓練を行った専門家が使用するため、字幕表示の遅延が短く、誤りも少ない。CART や C-Print は海外で使われている。文献[4]によると、海外でのリアルタイム字幕入力では速記者または要約筆記者 1 名で行う。日本では 4 名で行っている。4 名の内に入力は 2 名で行い、1 名が文の前半を入力し、1 名が文の後半を入力するように役割を分担する連係入力を行う。残り 2 名は交代のために待機している。文献[6]によると、パソコンの入力ウィンドウの上に設置されたモニターウィンドウに入力の様子がリアルタイムで表示される。そのため、連係入力を行う際に、片方の者がどこを打っているのか確認しながら入力できる。文献[7]によると、連係入力では、入力者間でどこまでの入力を担当するかを決める意思疎通が難しいとされている。これを解決するために、文節ごとの交代などを決める、合図する、モニター部でのパートナーの入力状況を監視するなどの方法が工夫されている。入力者同士の意思疎通は、入力作業と並行して行うため、片方の者が入力に手間取るか、入力ミスが発生すると連係が破綻してしまう。その場合、入力文の重複や欠落が発生し、修正作業を誘発してしまう。

2.2. SCRIBE

音声文字起こしシステムの課題を解決するために、クラウドソーシングを用いた音声文字起こしシステムである SCRIBE[8]が開発されている。このシステムでは、図 2.1 のように、まず、音声ストリームを自動で、一定時間の区間として短く区切る。次に、複数の非専門家のワーカで分担して区間の音声を文字起こしする。さらに、文の欠落や重複を無くすために、複数のワーカによって生成された前後の文を連結して字幕として受講者に提供する。文の連結にあたり区間の間を重複させておくことで、重複した文の単語順などにより統合され、重複や欠落のない字幕が生成される。

また、ワーカの負担を下げるために TimeWarp[9]も提案されている。これは、区間内の音声ストリームの再生速度を遅くし、遅延を取り戻すために区間外の音声ストリームの再生速度を速くする方法である。これは、ワーカの入力速度と講演者の発話速度に大きな差がある場合に有効である。また、区間内の再生速度を遅くすることで、入力終了時に遅延が発生し、文脈の把握が困難となる。それを防ぐために区間外は早送りで音声聞きながら遅延を取り戻す。このシステムは、聞き溜めや非常に速いタイピング速度などの技能を持たない複数の非専門家のワーカだけでもリアルタイム字幕を提供できる。しかし、文の欠落と重複を無くすために複数人の入力結果を統合するが、手話文字通訳は音声の文字起こしとは異なり、手話から日本語への通訳を要する。つまり、通訳によって表出された単語は同様の方法で統合できない場合がある。また、TimeWarp では区間内の再生

速度を遅くし、区間外の再生を速くしているが、手話でも再生速度を遅くすることや速くすることについて有効であるかは自明ではない。

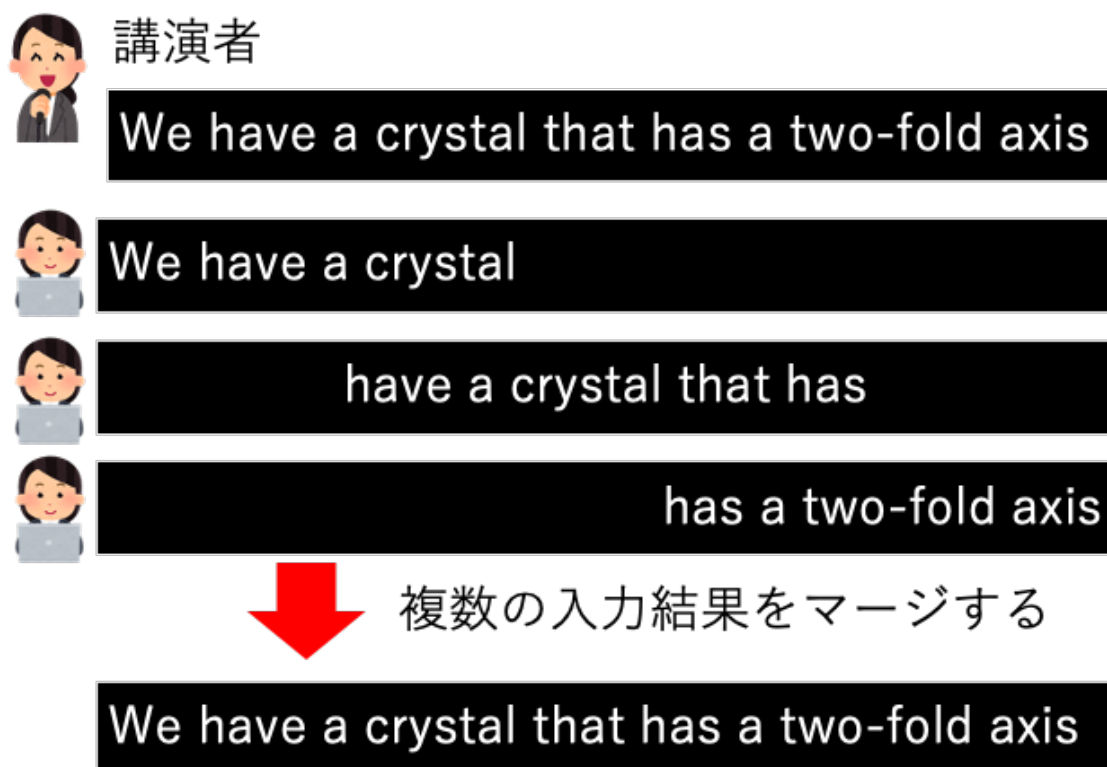


図 2.1 Scribe

2.3. クラウドソーシングを用いた手話文字通訳

クラウドソーシングを介した手話文字通訳に関する研究(以下、白石らの研究)[10]がある。この研究では、講演者のいる現場にワーカ全員が集まり、ワーカは講演者の手話を読み取りながら役割に合った作業を行う。流れを図 2.2 に示す。まず、講演者は手話で発話し、ワーカ全員は現場で講演者の手話を読み取り、役割に合った作業を行う。次に、通訳区間を作成する作業を担当する 1 名のワーカ(以下、カウンタ係)が意味単位で通訳区間を作成し、1 つの通訳区間は 1 グループ 3 名以上の通訳を担当するワーカ(以下、通訳ワーカ)に割り当てられる。さらに、グループ内の各通訳ワーカは、通訳区間を割り当てられた時、手話文字通訳を行い、図 2.3 のように複数の入力結果の文章が生成される。最後に、投票を担当するワーカ(以下、投票ワーカ)が 1 グループの入力結果から通訳品質の高い最適文を選び、字幕に表示する。手話文字通訳は音声文字起こしとは異なり、日本語への通訳を要するため、区間の間の単語の統合が困難である。カウンタ係が講演者の手話の意味の切れ目を見て、通訳区間を区切るため、通訳区間の間の文の欠落や重複が発生しにくい。しかし、通訳区間を意味単位で作成する方法では、通訳区間の時間が一定ではなく、テキストを入力するワーカの作業量や通訳品質はカウンタ係の能力に依存する。また、カウンタ係は

常に手話を見なければならないため、負担が大きい。1 グループ 3 名以上の理由は、白石らの文献によると、40%以上の通訳品質を保障するために、必要な 1 グループのワーカ数を推定した結果、3名以上であったからである。しかし、通訳品質を保障するために、1 グループ 3 名以上を 5 グループ以上の通訳ワーカと投票ワーカ、カウンタ係を必要とするため、人的コストが高い。

このシステムの拡大や改良に関する研究が 4 件報告されている。1 つ目の方法は、ワーカにできるだけストレスを与えないようにしながら、ワーカがグループに参加する時や離脱する時に、各グループ内の人数を均等にするように入れ替える方法である[11]。2 つ目の方法は、ワーカがグループに参加する時や離脱する時に、ワーカの能力のバランスを取りながらワーカを入れ替える方法である[12]。3 つ目の方法は、ワーカのグループが文章に通訳するタスクを実行した後、グループの通訳結果文を比較し、1 つ目の文に欠落している単語があり、2 つ目の文ではその単語が欠落していなければ、1 つ目の文に欠落していた単語を補完することで通訳品質を高める方法である[13]。4 つ目の方法は、ワーカに割り当てられるタスクの数をできるだけ均等にしながら、ワーカの能力に見合ったタスクを割り当てる方法である[14]。しかし、これらの研究では、人的コストの課題は解決されていない。

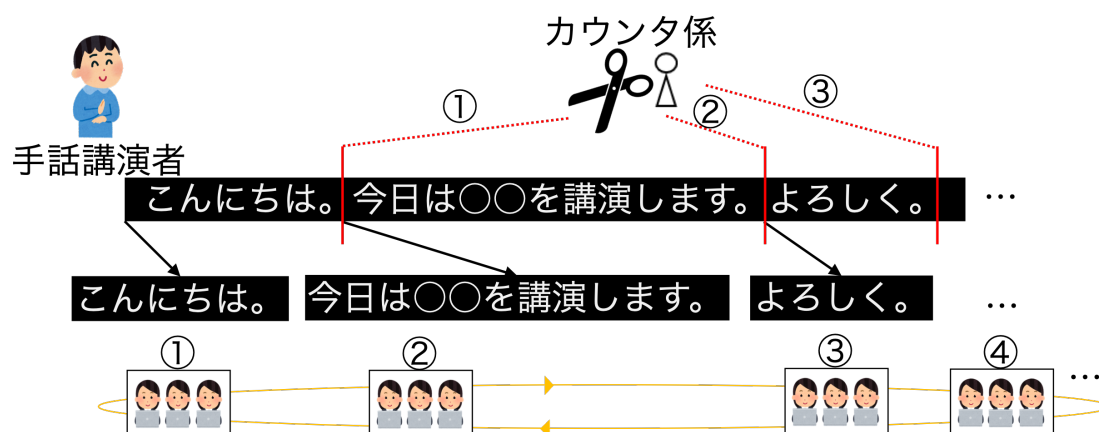


図 2.2 クラウドソーシングを用いた手話文字通訳システム

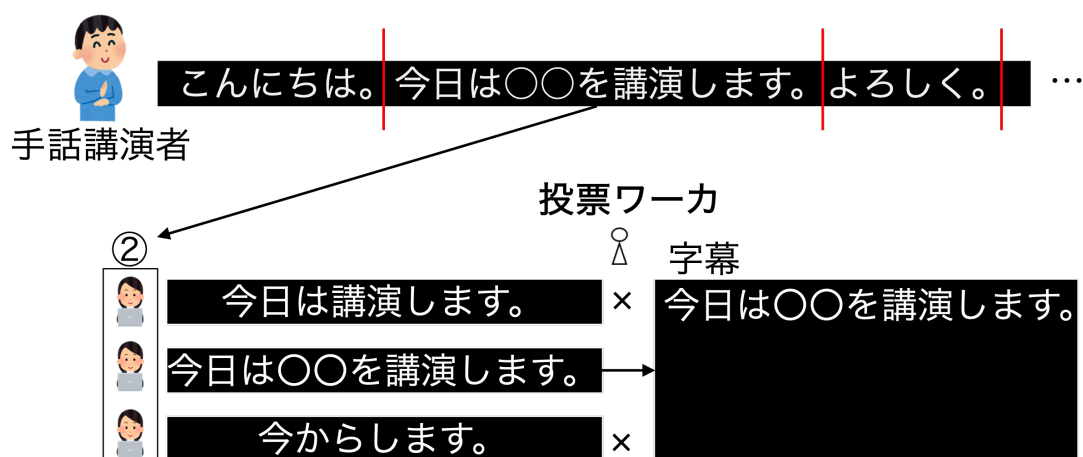


図 2.3 グループの入力結果の評価方法

2. 4. 本研究で実装するシステム

本研究では、人的コストの課題を解決するために、1つの通訳区間をグループではなく、1名のワーカが担当する方法を提案する。さらに、カウンタ係をなくし、字幕表示の遅延を減らすために、SCRIBEを参考にシステムが通訳区間を自動で、一定時間で区切り、通訳区間を重複する。手話文字通訳は音声文字起こしとは異なり、通訳を行うため、通訳区間の間の文の欠落や重複を無くするために前後文を統合する方法は適さない。その代わりに、ワーカインタフェースには前後の通訳区間を担当するワーカの入力の様子を表示し、連係入力が可能にする。本研究と従来の連係入力の違いは、文字通訳者2名ではなく、複数人のボランティアのワーカ同士で行うことである。そのため、複数人同士が文節の交代や合図などの意思疎通を行う方法は困難である。この課題を解決するために、通訳区間を自動で区切ることで、ワーカは入力担当範囲が分かるようにした。つまり、ワーカは意思疎通を行わなくても、前後のワーカの入力文を確認するだけで連係入力が可能になり、文の欠落や重複を防ぐことが期待できる。

通訳品質を保障するために、投票の方法ではなく、ライブ動画配信を行い、動画操作機能を設ける。TimeWarpを参考に、発話に追いつきながら入力を行うために、区間内の再生速度を下げ、区間外は再生速度を速くする方法を採用した。本研究では、音声ではなく手話のため、再生速度を速くした時、ワーカは速く感じるかの主観評価を行う。また、手話と文字を交互に見て入力する時、手話の見落としを防ぐために、本研究では巻き戻しや停止の動画を制御する機能を設けることを提案し、区間内の再生速度を下げる機能を含めてどのように使うか、主観評価を行う。

これらの方法により、白石らの研究と本研究と比較を行うために、通訳時間の客観評価や通訳品質の主観評価、必要人数の推定を行う。

第3章 手話文字通訳システム

3.1. システムの概要

本システムでは、白石らの研究の課題を解決するために、まず、講演者はライブ動画を配信する。次に、通訳区間を自動で、一定時間で短く区切り、複数人のワーカに割り当てる。さらに、タスク（担当通訳区間の手話文字通訳作業）が割り当てられた1名のワーカはクラウドソーシングのタスクとして、手話動画を見ながら手話文字通訳を行う。最後に、入力結果は通訳区間の順に連結して字幕に表示される。ワーカは、タスクを受け付けるためやタスク完了後にキューに入る。つまり、ワーカがタスクを早く完了するほど、より多くのタスクが割り当てられる。

字幕の遅れを減らすために、通訳区間は自動で、一定時間で区切り、各通訳区間は各ワーカにタスクとして割り当てる。しかし、通訳区間を一定時間で区切ると、通訳区間の中で手話が途切れることがあり、欠落が発生しやすい。欠落を防ぐために通訳区間を重複する。しかし、重複させても通訳区間の間の文の重複と欠落が発生してしまう[1]。これを解決するために、ワーカインタフェースには前後のワーカの入力文をリアルタイムで表示し、通訳区間の間の文の欠落や重複をなくす連続入力を検討する。

1つの通訳区間に対し、1名のワーカで通訳品質を保障するために、ワーカインタフェースに映像の見直しや再生速度を下げるなどの動画操作機能を設けた。しかし、区間内で動画操作機能を用いると、タスク完了後に遅延が発生する。タスク完了後にライブの再生位置に移動する方法も考えられるが、文脈の把握が困難となる。そこで、区間外では、ライブの再生位置に追いつくまでに早送りで手話を読み取る方法を検討する。

3.2. システムアーキテクチャ

システムアーキテクチャを図3.1に示す。ライブ動画配信にはYouTube Live[15]を用い、管理画面と通訳画面、字幕画面を設けた。管理画面には、自動でライブ動画の通訳区間を一定時間で区切り、ワーカにタスクを割り当てるためのタスク割り当て開始機能がある。通訳画面では、動画を見ながら手話文字通訳を行うためのユーザインタフェースを提供する。字幕画面では、各ワーカの入力結果をタスクの順番に表示する。サーバーはNode.js[16]で実装した。クライアントはWebサイト経由で作成し、YouTube Player API[17]を利用して、Webサイトで動画の閲覧や制御ができるようにした。

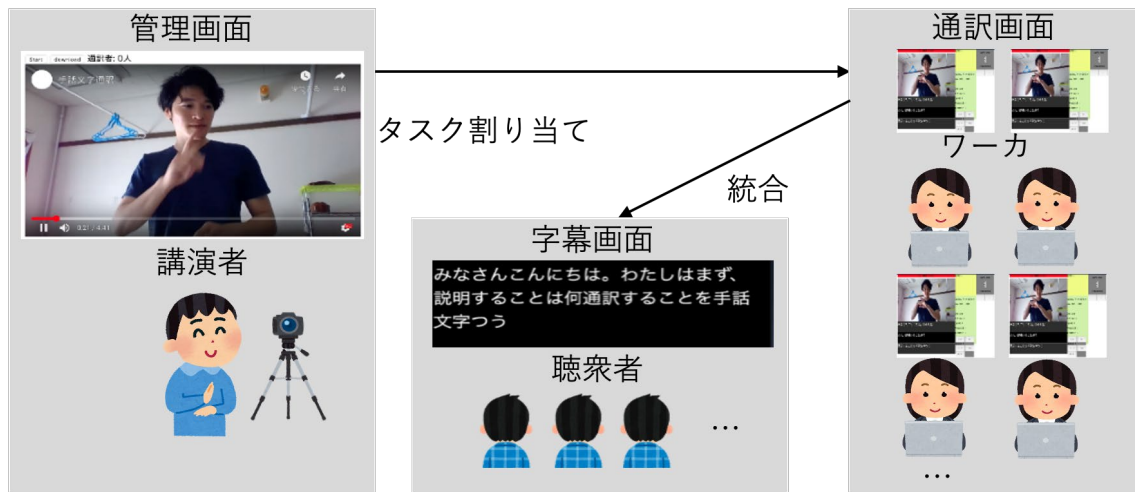


図 3.1 手話文字通訳システムのアーキテクチャ

3.3. 管理画面

管理画面は通訳区間を自動で区切り、手話文字通訳のタスクとしてワーカに割り当てるためインタフェースである。図 3.2 のように、まず、講演者が管理画面の「start」ボタンをクリックする。次に、ライブ動画の通訳区間を自動で、一定時間で区切ると共に通訳画面を用いるワーカにタスクを割り当てる。さらに、タスクはワーカキュー内のワーカに割り当て、ワーカキューから離れると共に手話文字通訳を開始する。最後に、タスクを完了したワーカはワーカキューに入る。

YouTube Live から配信される動画は、同一動画を見ているクライアント間のディレイは異なる。「start」ボタンをクリックした時、管理画面の動画の再生位置を基準に、サーバーに接続されている全ての通訳画面の動画の再生位置を同期させた。

「download」ボタンを押すと、ログデータのファイルをダウンロードできる。ログデータは通訳画面でのワーカがキーボード入力した時のタイムスタンプや入力した文字などがあり、これらはサーバーを介して管理画面に送信する。

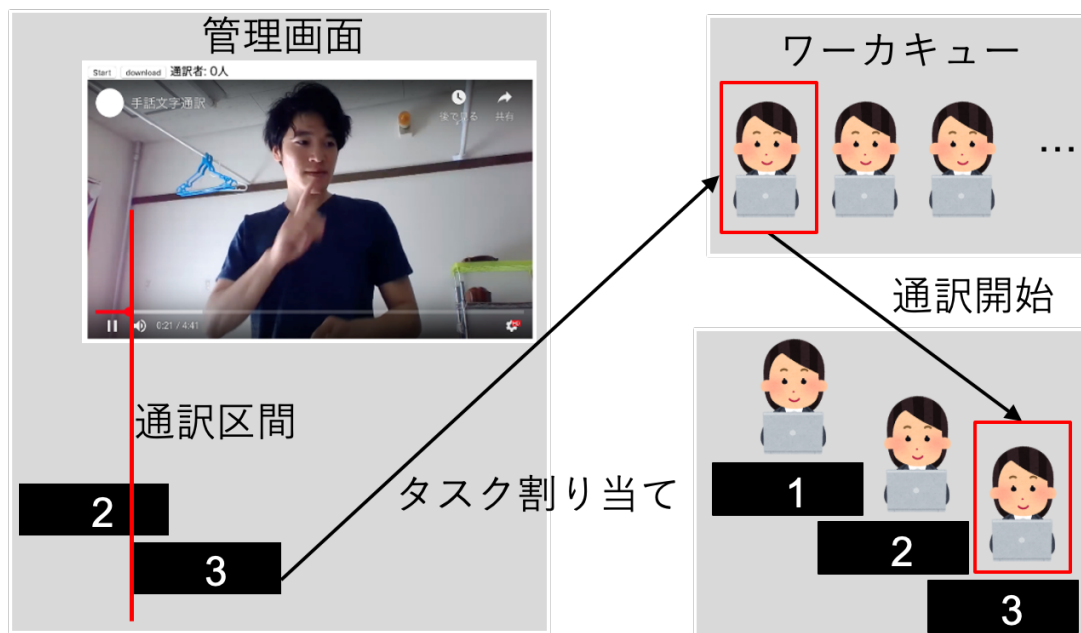


図 3.2 管理画面

3. 4. 通訳画面

3. 4. 1. 通訳画面の動画と入力枠

4章の予備実験で使用した通訳画面を図 3.3 に示す。通訳画面はワーカの手話文字通訳を容易にするためのユーザーインターフェースを提供する。ワーカは手話と文字を交互に見ながら入力するため、視線移動を減らすために動画の下に、入力枠を表示した。タスクが割り当てられたワーカは、通訳区間内の手話動画を見ながら、入力枠に文字を入力する。

3. 4. 2. 動画操作機能と早送り

ワーカは手話の見落としを防ぐために、動画操作機能として以下の 3 つがある。

- ・「1 秒戻る」
- ・「再生・停止」
- ・「再生速度調整」

「1 秒戻る」とは、動画の現在の再生位置の 1 秒前に巻き戻すことである。連打や長押しを行うことで、より速く巻き戻すことができる。手話を見落としした時、または読み取れなかった時に巻き戻すことで再確認が可能である。「再生・停止」とは、動画の再生と停止の切り替えを行う機能である。ワーカは手話を覚えられる量に限りがあり、忘れる前に動画を停止して文字入力することができる。「再生速度調整」とは、タスクが割り当てられた時、区間内のみワーカが指定した再生速度に変更できる機能である。再生速度は 0.1 倍速の間隔で 0.3 倍速から 1.0 倍速の間とした。ワーカが発話速度とタイピング速度に差がある場合、再生速度を下げることで発話速度とタイピング速度の差を縮め

ることができる。これらの動画操作機能はショートカットキーとボタンによって操作できる。ショートカットキーの操作方法は、Windows では Alt キーの同時押し、macOS では Ctrl の同時押しとした。両方同じにしなかった理由は、Google Chrome ウェブブラウザ[17]にもショートカットキーが用意されており、通訳画面のショートカットキーと混合することを防ぐためである。

ワーカが区間内で動画操作機能を用いると、タスク完了後の再生位置とライブの再生位置の間に遅延が発生してしまう。タスク完了後、ライブの再生位置に同期する方法も考えられるが、文脈の把握が困難となってしまう。解決するために、区間外では、ライブの再生位置に追いつくまで再生速度を上げることにした。つまり、早送りで遅延を取り戻しながら手話動画を視聴することになる。

3.4.3. 指示枠

動画の上に現在の状況をワーカに伝達するために、ワーカへの指示文が表示される領域(以下、指示枠)がある。指示文には以下の 5 種類があり、自動で表示する。

- ・「まだ始まっていません」
- ・「通訳中」
- ・「リアルタイムに追いつくための早送り」
- ・「待機中 後○人です」
- ・「次はあなたの番です」

「まだ始まっていません」とは、タスク割り当てを開始してない時に表示される指示文である。具体的には、管理画面のスタートボタンを押し始めるまではこの指示文が常に表示される。

「通訳中」とは、タスクが割り当てられた時からタスクを完了するまでの間に表示され、その間にワーカは手話文字通訳を行う指示文である。「通訳中」の時は手話文字通訳を終えるまで通訳区間の外へ出られないようにしている。つまり、通訳区間の前後の再生位置に移動した場合、通訳区間の初めや終わりの再生位置に戻すようにしてある。

「リアルタイムに追いつくための早送り」とは、タスク完了時とライブの再生位置の間に遅延が発生していた場合、早送りで遅延を取り戻す指示文である。ライブの再生位置に追いつくまでの時間も表示される。

「待機中 後○人です」とは、次のタスクが割り当てられるまで通常の再生速度で動画を見ながら待機する指示文であり、○の部分に通訳開始までの人数が表示される。この指示文は「リアルタイムに追いつくための早送り」で遅延を取り戻した時や、タスク完了後に遅延がなかった時に表示する。

「次はあなたの番です」とは、待機中のワーカが次のタスクを担当とする場合に表示され、通訳の準備を始めることができる。

例外として、ワーカ的人数が足りない場合、「リアルタイムに追いつくための早送り」の時、ワーカが誰も担当していない通訳区間の初めの再生位置に追いついた時にタスクを割り当てる。つまり、遅延なく実現するためのワーカ人数が不足している時、ライブの再生位置に追いつかない限り区間外では「待機中 後○人です」になることなく常に早送りとなる。

3.4.4. 前後入力文

通訳区間の間の文の欠落をなくすために、前後で1秒間重複させた。しかし、通訳区間を重複しても文の欠落と重複が発生してしまう。そこで、ワーカの入力枠の上下に、前後の通訳区間の入力枠(以下、前後入力文)を表示した。前後のワーカの入力中の文はリアルタイムに表示される。これでワーカは入力中に前後の入力文を確認することができる。この時、文の欠落や重複に気づいたワーカが連係入力を行うことで、欠落や重複を無くせるか検討する。また、入力中に前後の入力文を確認することで、自分の入力文の文脈に誤りがないか確認することができる。

3.4.5. 動画サイズ, 文字の大きさ, ワーカの人数

パソコンによって動画サイズや文字の大きさが異なるため、動画サイズや文字の大きさを調整できるようにした。動画サイズは「大=800px:450px」, 「中=640px:360px」, 「小=480px:270px」とした。文字の大きさは 1px 間隔で変更でき, 「↑」のボタンを押すと 1px 大きくなり, 「↓」のボタンを押すと 1px 小さくなる。

ワーカが現在何人参加しているか確認できるようにするために、通訳者の人数を表示させた。



図 3.3 通訳画面

3.5. システムの不具合対処方法

本システムでは、システムの不具合が発生しないように、研究室の学生や先生に試してもらった。しかし、本番では使用機種や通信方法によって接続タイムアウトや画面が固まるなどのシステムの

不具合が発生する可能性がある。現状のシステムでは、不具合が発生した場合、通訳画面の Web サイトを更新するしか解決策がない。つまり、不具合が発生したワーカは途中で抜けるは必然である。途中で抜けたワーカは、ライブの再生位置から開始する方法も考えられるが、それまでの手話を読み取ることができず、講演内容の理解が困難となる。そこで、本システムでは、ワーカが途中で抜けた時の再生位置を記憶し、戻った時にその再生位置から開始できるようにした。つまり、接続したとき、抜けた再生位置から開始し、遅延を取り戻した後、手話文字通訳に参加できる。

3.6. 字幕画面

字幕画面には、複数のワーカが通訳画面で手話文字通訳を行った結果を、図 3.4 に示す。通訳区間の順に連結して字幕として表示する。ワーカの入力途中の文もリアルタイムで表示されるようにした。本論文では、字幕の最適な表示方法について着目していないため、今後の課題とする。

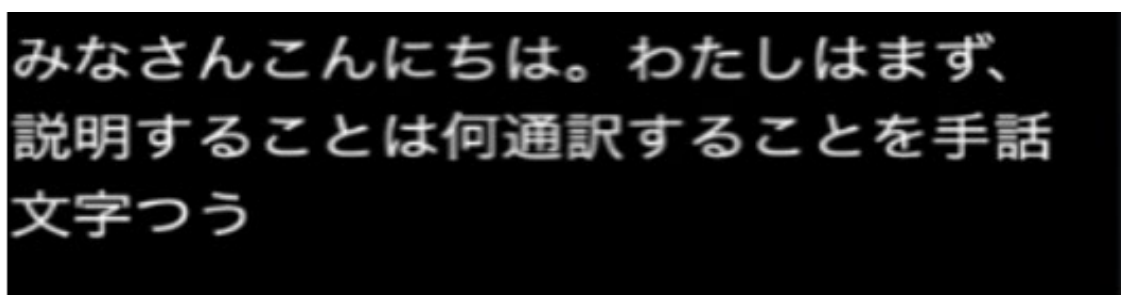


図 3.4 字幕画面

3.7. システム全体の処理

手話文字通訳システム全体の処理の流れを図 3.5 に示す。最初に、ワーカが通訳画面を開いたら、通訳画面(以後、ここではクライアント)から管理画面(以後、ここではサーバ)に参加を要求する。要求を受けたサーバでは、ユニークなワーカ ID を設定しクライアントに返送する。また、クライアントに全クライアント数を送信し表示させる(第 3.4.5 節)。

講演者かシステム管理者が管理画面上の start ボタンを押したら、通訳画面を開いている全ワーカの ID をキューに挿入すると共に、サーバの動画の再生位置とクライアントの動画の再生位置を同期させる。次に、サーバでは、講演終了までに通訳区間を一定時間で、自動で区切る(第 3.3 節)。区切ると同時に、キューの確認を行い、キューの 3 番目以降のクライアントには、そのキューの位置に応じて「待機中 後〇人です」の指示文を送り表示させる。キューの 2 番目のクライアントには、「次はあなたの番です」の指示文を送り表示させる。後 0 人、つまり先頭のクライアントには、「通訳中」の指示文を送り表示させ、そのクライアントのワーカは入力を開始する(第 3.4.3 節)。これと同時に当該クライアントをキューの先頭から取り出す。また、自動でワーカが指定した通訳時の再生速度に変更される(第 3.4.2 節)。入力途中文や入力結果を字幕画面に送り表示させる。この際、入力途中文は他の通訳中のクライアントにも送られ、担当する通訳区間の前後関係に従って、前／後の通訳者の入力文の領域に表示させる(第 3.4.4 節)。入力終了したら、サーバに入力終了を

送信する。サーバが入力終了を受け取ったら当該クライアントの再生位置を考慮しながらキューへ挿入しなおす。

入力終了したクライアントでは、クライアントの再生位置とサーバの再生位置と同期しているかについて、サーバに対して現在の再生位置の情報とともに確認の要求を行う。なお、サーバでは再生位置の差が 1 秒未満であれば同期しているものとみなすように設定した。同期していなかった場合、クライアントの動画では早送りの再生速度に変更され、サーバから「リアルタイムに追いつくための早送り」の指示文を送り表示させる。その 1 秒後に、再びクライアントから同期したかの確認の要求を行う。同期した場合、クライアントの動画の再生速度が 1.0 倍速に変更される。その後は、「待機中 後〇人です」または、「次はあなたの番です」の指示文が表示され、通訳が開始するまで待機することになる。クライアントは動画を最後まで再生し終えたら完了となる(第 3.4.3 節)。

図 3.5 では、本システムで理想的に手話文字通訳がなされている状況について述べた。しかし、遅延なく実現するためのワーカの人数、つまりクライアント数が足りていなかった場合、誰も担当していない通訳区間が残ってしまう。この場合、サーバの再生位置ではなく、誰も担当していない通訳区間の開始位置に、通訳中ではないクライアントの動画の再生位置が達した時に、そのクライアントに当該通訳区間を割り当てる。システムの不具合などが発生したクライアントは途中で離脱する。この状況を想定して、サーバとクライアントが接続して以降、2 秒ごとに接続確認を行う。接続確認とともに、クライアントに現在の再生位置を送信させ、サーバでその再生位置をクライアントごとに保存、更新する。そのため、あるクライアントが途中で離脱したとしても、クライアントごとの最新の再生位置が保存されているため、離脱したワーカが戻った時に離脱時の再生位置を送ることができ動画の見落としを防ぐことができる(第 3.5 節)。

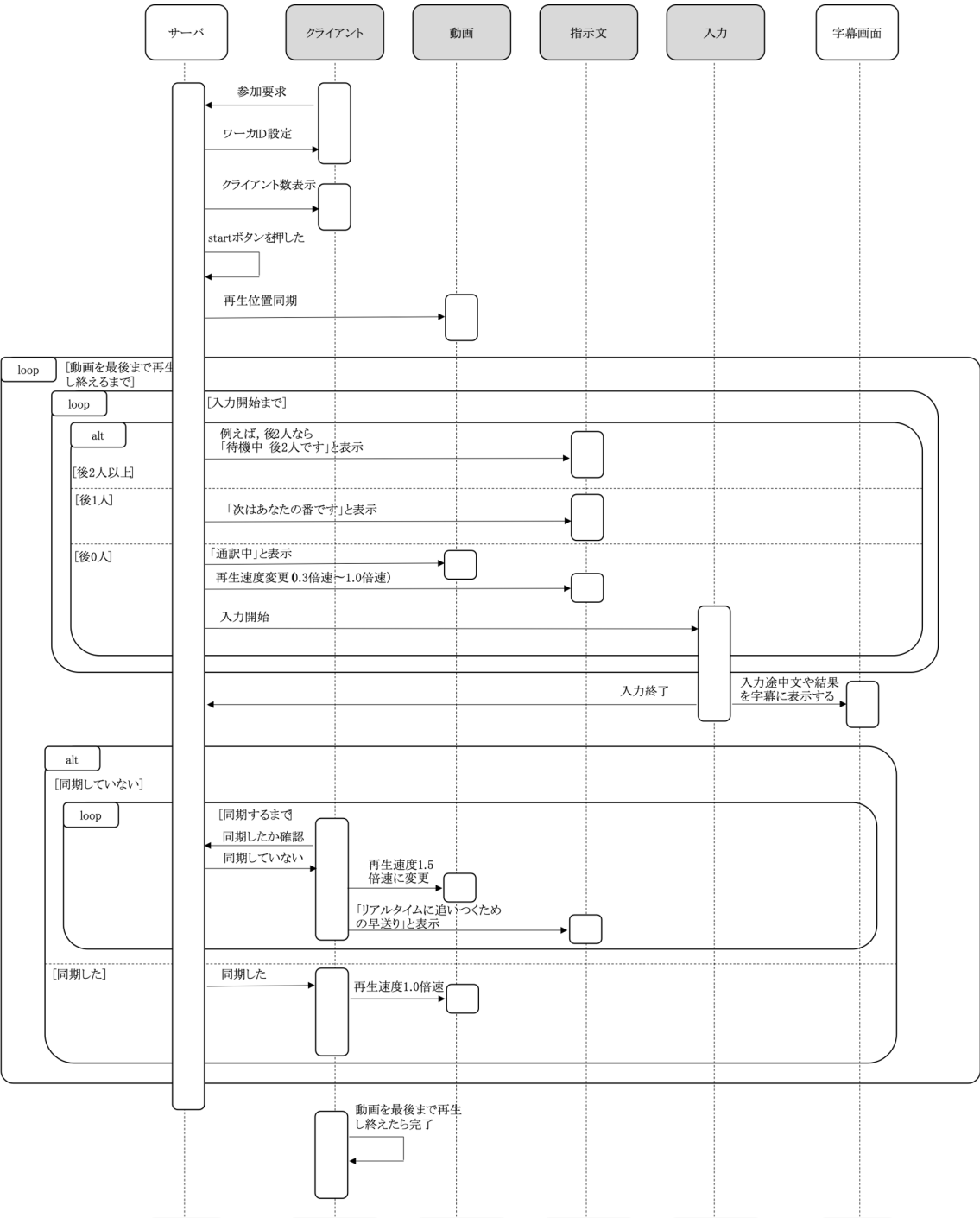


図 3.5 手話文字通訳システムの処理の流れ

第4章 手話文字通訳の予備実験

4.1. 実験目的

プロトタイプシステムを用いた手話文字通訳を通して、以下の評価や調査を行う。

- ・本システムでの通訳時間の客観評価
- ・通訳区間の間の文の欠落率と重複率の客観評価
- ・遅延を取り戻すための早送り再生速度 1.5 倍速についてどう感じたかの主観評価
- ・通訳の楽しさの主観評価

本システムでは、字幕の遅れを減らすことができたか確認するために、タスク(担当通訳区間の手話文字通訳作業)開始から入力終了までの時間(以下、通訳時間)の客観評価を行う。通訳区間を自動で、一定時間で区切り、重複すると、文の欠落や重複が発生するため、連係入力で解決できるか客観評価を行う。遅延を取り戻すための早送りについて、手話の場合でも問題ないか明らかにするために、主観評価を行う。ボランティアが本システムを継続利用してもらうには、通訳の楽しさが重要だと考え、主観評価を行う。

4.2. 実験環境と条件

実験の流れは、まず、実験に参加するワーカーに日本語入力スピード測定[18]を用いてタイピング速度を記録する。次に、本システムについて説明を行い、本システムに慣れてもらうために訓練用動画を見て手話文字通訳を1回行ってもらい。さらに、実験用動画を見ながら手話文字通訳を行ってもらい、ログデータを収集する。最後に、アンケートを回答してもらった。本システムについて説明したことは3つある。1つ目は、通訳区間の間の文の欠落と重複、連係入力について説明し、できるだけ連係入力を行って欲しいと伝えた。2つ目は、動画操作機能について説明し、発話が早く感じるなら再生速度を下げるか停止、手話の見落としがあれば1秒巻き戻しを使っていと説明した。3つ目は、手話文字通訳を行う時、分からない手話があれば諦めていいことや、リラックスして作業してほしいと伝えた。理由は、速く入力してほしい、正しく通訳してほしいと伝えると、プレッシャーを感じ、楽しめなくなると考えたからである。

実験参加者であるワーカーはろう者・難聴者の男子大学生4名であり、手話歴は全員7年以上であった。ワーカーには各自が所持しているノート型パソコンを使用してもらい、システムの不具合があった2名にはノート型パソコンを貸与し行ってもらった。管理画面や通訳画面の表示と操作にはGoogle Chrome ウェブブラウザ[19]を利用した。本システムはYouTube Liveに対応しており、ライブ動画に対する手話文字通訳が可能であるが、今回の実験では、事前に講演者を撮影して保存した手話動画をYouTubeにアップロードして利用した。実験を実施した日程は2019年10月18日である。訓練用動画の長さは4分40秒であった。実験用動画の長さは6分24秒であった。講演者は日本語対応手話と日本手話が混成した中間手話であり、口形を使用していた。手話の速さの評価方法は、動画を見て、各通訳区間9秒間の内に表出された手話単語数をカウントした。通訳

区間の終わりに途切れた手話もあったが、それも含めてカウントした。理由は、途切れた手話でも読み取れた場合、ワーカはその手話を通訳することが想定できるからである。指文字は、例えば指文字を5回表出した場合、5単語としてカウントした。助詞は、指文字で表した場合のみカウントし、口形のみであればカウントしなかった。結果、手話の速さは1秒で1.5単語、通訳区間9秒の時13.4単語であった。NHK週間手話ニュースの映像[20]では、手話の速さは1秒で約1.6単語であり、少し遅い程度であった。通訳区間は48区間であった。通訳区間の長さは9秒とした。理由は、白石らの研究[10]の通訳区間の平均区切り時間を参考にしたからである。通訳区間の重複時間は1秒とした。理由は、我々のこれまでの研究[1]を参考にしたからである。タスク完了後の遅延を取り戻すための再生速度は1.5倍速とした。理由はTimeWarp[9]を参考にしたからである。

4.3. 実験結果

4.3.1. 各ワーカのタイピング速度と通訳時間

各ワーカのタイピング速度、通訳時間の平均と標準偏差を表4.1に示す。タイピング速度とは1秒にひらがなで入力された数値の平均であり、ワーカのタイピング速度の平均は、1.9CPS (characters per second) (SD 0.3CPS)であった。ワーカがタスク開始から終了までかかった時間(以下、通訳時間)の平均は25.9秒(SD 13.2秒)であった。各ワーカの通訳時間とタイピング速度の散布図を図4.1に示す。タイピング速度と通訳時間を、ピアソンの積率相関係数を用いて相関係数を求めたところ、中程度の負の相関が認められた($r = -0.68$)。よって、タイピング速度が高いほど通訳時間が短くなる傾向があるといえる。

表 4.1 タイピング速度と通訳時間の平均と標準偏差

ワーカ	タイピング速度[CPS]	通訳時間の平均[秒]	通訳時間の標準偏差[秒]
A	2.4	16.7	2.2
B	1.9	25.7	9.9
C	1.7	45.2	8.3
D	1.6	33.6	14.7
全員	1.9	25.9	13.2

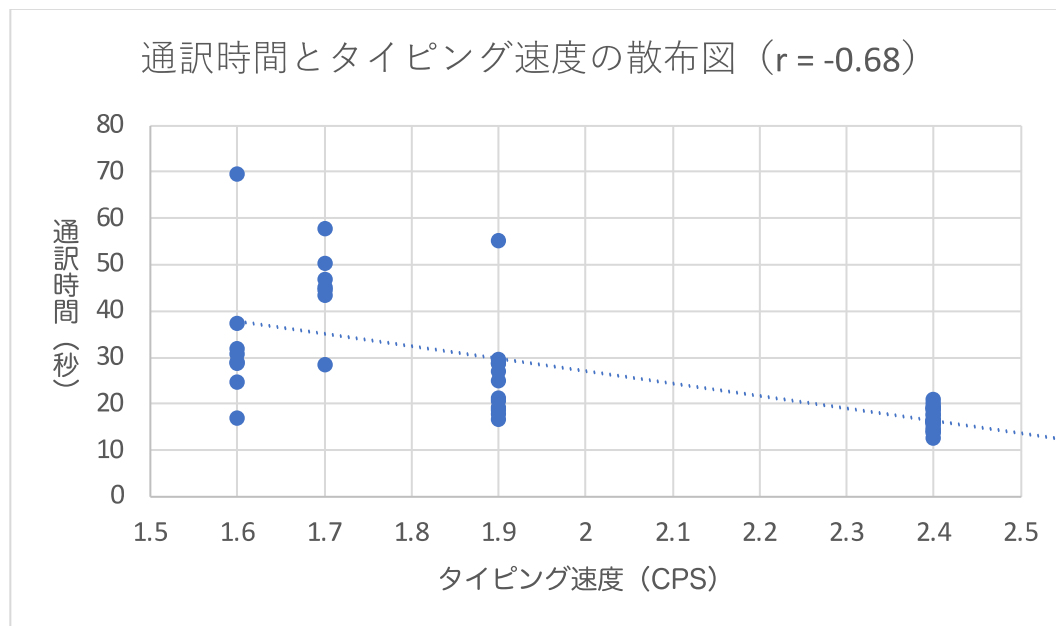


図 4.1 通訳時間とタイピング速度の散布図

4. 3. 2. 通訳区間の間の文の欠落と重複のある通訳区間の割合

通訳区間の間の文の欠落と重複のある通訳区間の割合を図 4.2 に示す. 欠落や重複のある通訳区間の割合を横棒グラフで示した. 欠落や重複のある通訳区間の調査方法として, 通訳区間の間にワーカの入力結果と原稿, 手話動画を見比べてカウントした. 通訳区間の間の文の重複の例は, 現在の通訳区間が「テレビで初めて」の時, 次の通訳区間が「初めて野球を見ました。」であれば, 「初めて」が同じ単語であり重複している. また, 「左打ちに変更」と「左利きに転向」の場合, 両方の通訳結果は異なるが, 手話単語は同じであるため, 重複である. 通訳区間の間の文の欠落の例は, 「きっかけが」と「よりもかっこいい」の場合, 通訳区間の間に「プレー」が入力されていないため, 欠落といえる. また, 時々通訳区間の間に助詞や句読点が入力されていない場合もあり, これも欠落といえる. 通訳区間は 48 区間あるが, 最後の通訳区間は次の通訳区間がないため, 省くことで 47 区間となった. その結果, 欠落と重複のある通訳区間の割合は両方 34%であった.

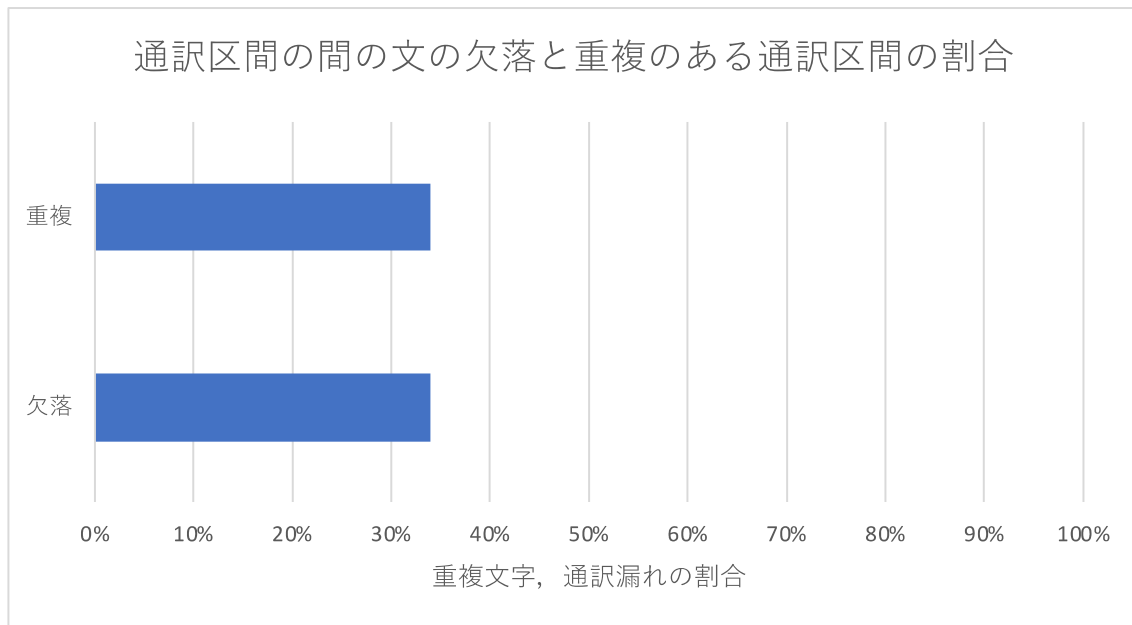


図 4.2 通訳区間の間の文の欠落や重複のある通訳区間の割合

4.3.3. 各ワーカが担当した通訳区間

各ワーカが担当した通訳区間を調査した結果を図 4.3 に示す。通訳区間は 48 区間あり，各ワーカが担当した区間を折れ線グラフで示した。その結果，通訳回数はワーカ A が最大で 21 回，ワーカ C が最低で 7 回であった。ワーカ A のみ 2 連続通訳を行っていたことが分かった。

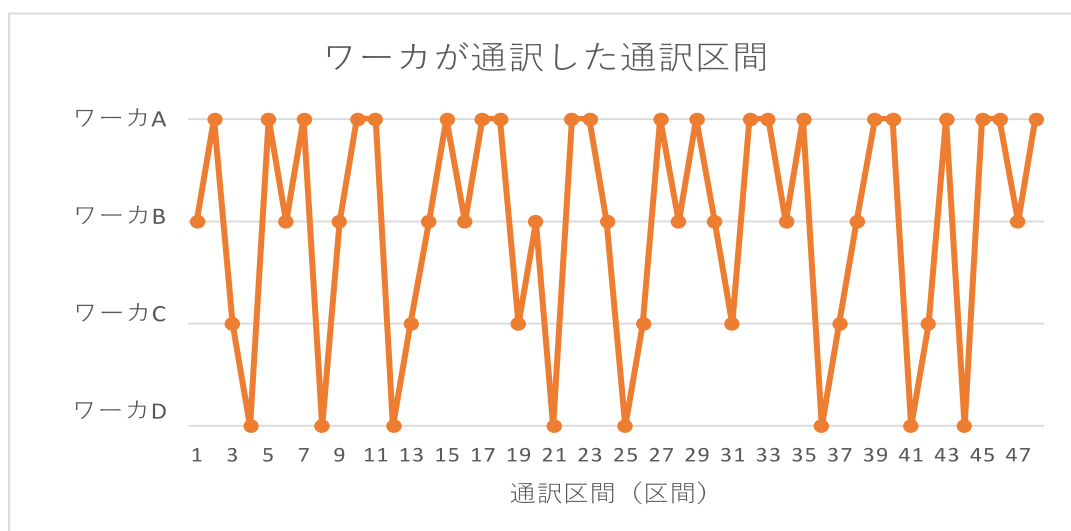


図 4.3 ワーカが通訳した通訳区間

4.3.4. 主観評価結果

アンケートで「通訳中に前後の入力文を確認しましたか(以下, 前後文を確認したか)」を2段階評価で「はい」と「いいえ」で回答してもらった。「リアルタイムに追いつくための再生速度 1.5 倍速は速く感じたか? (以下, 早送りは速く感じたか)」を5段階評価で「1=遅い」, 「3=ちょうど良い」, 「5=速い」で回答してもらった。「通訳は楽しかったですか」を5段階評価で「1=ネガティブ」, 「5=ポジティブ」と回答してもらった。これらの回答結果を表 4.2 に示す。

前後文を確認したかについて, 確認したワーカは1名であった。確認した者の回答理由は「文字の繋ぎでおかしい所がないよう確認した。」であった。確認しなかった者の回答理由は「自分の分を入力するだけで精一杯だったから(「いいえ」と回答)」であった。

早送りは速く感じたかについて, 全員が「3=ちょうど良い」と答えていた。回答理由は, 「話を読み取るだけなら多少早送りでも内容はわかるから」と「あまり違和感がなかった」であった。

通訳の楽しさは「2=やや楽しくなかった」が1人, 「3=どちらでもない」が2人, 「5=やや楽しかった」が1人であった。やや楽しくなかったワーカの回答理由は, 「ひたすら集中して入力するのは疲れるから」であった。どちらでもないワーカの回答理由は, 「慣れないから疲れしました」であった。やや楽しかったワーカの回答理由は, 「休憩できるところもあるのでゆっくりとできた。」であった。

表 4.2 各ワーカの主観評価結果

ワーカ	前後文を確認したか[Y/N]	早送りは速く感じたか	通訳の楽しさ
A	No	3	2
B	No	3	3
C	No	3	3
D	Yes	3	4

4.4. 考察

4.4.1. タイピング速度と通訳時間

通訳時間の平均は9秒の通訳区間で26.0秒(SD 13.1秒)であり, 約3倍の時間がかかっていた。通訳時間の標準偏差を見ても, 13.1秒とばらつきが大きく, 通訳時間の平均が一番早かったワーカ A では16.7秒(SD 2.2秒), 一番遅かったワーカ C では45.2秒(SD 8.3秒)のため, 約2~5倍の時間がかかっていた。白石らの研究[10]での中間手話の講演者に対して, 通訳時間は19.3秒であり, ワーカ A 以外は遅かった。各ワーカのタイピング速度と通訳時間には中程度の負の相関があったため, タイピング速度が速いほど通訳時間が短くなる傾向といえることがわかった。つまり, 通訳区間9秒は, タイピング速度が一番速かったワーカ A 以外にとって長かったことが考えられる。字幕の遅れが大きいほど, 受講者は講演や周囲の状況把握が困難となるため, できるだけ減らす必要がある。次章の実験では, 通訳区間を短くすることで, 白石らの研究より通訳時間が短くなるか検討する。

本研究では, 非専門家のワーカを対象としているため, 各ワーカの能力の個人差の問題を解消するために, 入力が遅いワーカには通訳区間の時間を短く, 入力が高いワーカには通訳区間の時

間を長くするといった方法も考えられる。しかし、ワーカ自身が通訳区間の時間を手動で調整する方法では、最適な時間に調整するのに時間がかかることが考えられる。今後の課題として、各ワーカの能力に合った通訳区間の時間に動的に調整する方法を検討する。

4.4.2. 連係入力

通訳区間の間の文の欠落と重複を減らすために、通訳画面に前後入力文を確認できる機能を設け、連係入力で解決できるか検討した。その結果、我々のこれまでの研究[1]での通訳区間の間の文の重複率は 56%であったものが、今回は 34%に減少した。理由として、前後入力文を確認したワーカは 1 名おり、前後の通訳区間の間の文を確認していたとコメントがあり、より自然に前後の文を繋げることを意識していたことが考えられる。しかし、手話文字通訳に専念することで連係入力を行う余裕がなかった者がいたため、通訳区間の間の文の欠落率や重複率は 34%であり、依然として残っていた。連係入力を行う余裕がなかったと考えられる原因として、今回の実験では、訓練 1 回のみであったため、訓練不足が原因で手話文字通訳や連係入力に慣れてなかったことが考えられる。次章の実験では、手話文字通訳システムに慣れてもらうために訓練回数を増やし、手話文字通訳と連係入力の慣れの主観評価を行う。さらに、訓練後、実験用動画の手話文字通訳を行ってもらい、通訳区間の間の文の欠落率と重複率が減少できたか客観評価を行う。

4.4.3. 遅延を取り戻すための早送り

タスクを完了した後、遅延を取り戻すために動画の再生速度を 1.5 倍速にする方法を、手話動画にも適用できるか検討した。その結果、実験に参加した全員のワーカにとって、主観的にちょうど良い速度であることが分かった。手話ニュース映像の好ましい再生速度を検討した報告[18]では、手話習熟度や手話コンテンツにより読み取りやすい再生速度が異なるとされている。つまり、ワーカの手話習熟度や手話講演者によって好ましい再生速度が異なることが考えられる。次章では、複数の手話講演者に対して、複数のワーカは再生速度 1.5 倍速が好ましいかについて主観評価を行う。

4.4.4. 通訳の楽しさと必要人数

ワーカは 4 名のみであったため、通訳時間が短いワーカはキューに入れられた後、すぐ新しいタスクを割り当てられた。そのため、通訳回数が多いワーカと少ないワーカの間では 3 倍の差があった。通訳の楽しさと通訳回数を分析すると、やや楽しくなかったと答えたワーカ A は通訳回数が一番多く、2連続で通訳する状況が発生していた。つまり、常に入力するような状況であったため楽しくなかったことが理由だと考えられる。やや楽しかったと答えたワーカ D は通訳回数が2番目に少なく、コメントでは休憩できる場所があったと回答していた。ようするに、ワーカ A より余裕があったことが理由だと考えられる。

ボランティアが利用することを想定すると、ワーカ A のような状況が発生すると、楽しめず、参加意欲が低下する。ワーカ全員がワーカ D のように、タスク完了後に休憩できる時間があれば、より通訳を楽しめる可能性がある。そこで、作業負担軽減と共にワーカ全員が楽しめるように、必要人数

Nw を推定する. Nw は, 通訳時間 Tt と再生位置同期に要する時間 Tc の合計を通訳区間の時間 Ts で割ることによって求まる.

$$Nw = \frac{Tt + Tc}{Ts} \quad (1)$$

ここで, Rc は早送りの再生倍速, Rs は標準再生倍速とすると, Tc は以下の式で求まる.

$$Tc = \frac{Tt - Ts}{Rc - Rs} \quad (2)$$

通訳時間が一番遅かったワーカを基に計算すると, $Nw = 14.95$ ($Tt = 45.2$, $Ts = 8s$ (1秒重複のため), $Rc = 1.5$ 倍速) で 15 名のワーカがいれば, 作業負荷軽減と共に, 楽しさを増やすことができる可能性がある. しかし, 本研究では非専門家のワーカを対象とするため, ワーカの楽しさだけではなく, 通訳画面の操作のしやすさや理解のしやすさなどのユーザビリティも評価する必要がある. 次章の実験では, 必要人数に近いワーカ数で実験を行い, 通訳の楽しさと SUS の主観評価を検討する.

4. 5. まとめ

本章では, クラウドソーシングを用いて手話の読めるワーカが手話文字通訳を行うためのプロトタイプシステムを開発し, ワーカ4名で手話文字通訳の予備実験を行った結果を述べた. 分かったことは 4 つある. 1 つ目は, 字幕の遅れを減らすために通訳区間を9秒としたが, 白石らの研究より平均通訳時間が遅かった. 2 つ目は, 連係入力を行う方法で通訳区間の間の文の欠落や重複を減少できるか検討した. 結果, 連係入力を行ったワーカが1名いたため, 通訳区間の間の文の重複率が 56% から 33% に減少した. しかし, 連係入力を行う余裕がない者もいたため, 文の欠落や重複が依然として残っていた. 3 つ目は, タスク完了後の遅延を取り戻すための再生速度 1.5 倍速が速く感じるか主観評価を行った結果, 全員がちょうど良いと回答した. 4 つ目は, 通訳時間が一番遅かったワーカを基に必要な人数を推定した結果, 15 名のワーカが必要であった. この人数で行えば作業負荷軽減と共に通訳の楽しさが向上する可能性がある.

予備実験で明らかにした課題は, 3 つある. 1 つ目は, 訓練不足であれば, 手話文字通訳に精一杯で連係入力を行う余裕がないことである. 2 つ目は, 通訳区間の時間9秒では通訳時間の平均値が長く, 字幕の遅れが大きかったことである. 3 つ目は, 予備実験でのワーカ人数4名であり, 推定した必要人数 15 名より少なかったことである. さらに, 入力の速いワーカの負担が大きくなり楽しめなかった. 次章の実験では, 本章の結果を踏まえて, 通訳区間の長さを短くし, 必要人数に近いワーカ数で行う. また, 実験用動画の手話文字通訳を行う前に, 本システムに慣れるために訓練回数を増やす. そして以下の評価を行う.

- 通訳品質の主観評価
- 通訳区間を短くした時の通訳時間の客観評価と必要人数推定

- 手話文字通訳や連係入力、動画操作の慣れの主観評価
- 通訳区間の間の文の欠落率や重複率の客観評価
- 通訳の楽しさや SUS の主観評価
- 複数人の講演者に対して、複数人のワーカが再生速度 1.5 倍速は速く感じるかの主観評価

これらの評価を行い、提案するシステムの有効性を明らかにする.

第5章 手話文字通訳の本実験

5.1. 改良した通訳画面

本章の実験を行う前に、同研究室の学生や指導教員に通訳画面を使用してもらい、頂いた意見を基に、図 5.1 の通りに通訳画面を改良した。変更点のみ述べる。指示枠は動画の下に表示させた。指示文に「通訳中」が表示されている時、通訳区間の現在位置を表示するためのプログレスバーを指示文の下に表示させた。これによって、通訳区間の範囲や区間内の現在位置が直感的に分かるようになっている。「通訳中」の時に、通訳区間の終わりに到達したら、自動で通訳区間の初めに巻き戻されるようにした。つまり、通訳区間の初めに戻ることによって、見落とした手話の確認や手話と入力文の比較などが、手動で巻き戻すより容易にできるようになっている。動画操作機能を使用できるボタンは一番下に表示させた。動画操作機能である「通訳中の再生速度」はデフォルトで 0.8 倍速とした。1.0 倍速がデフォルトであると、発話速度とタイピング速度の差が大きくなるワークが多く、0.5 倍速では遅く感じると意見が多かったためである。動画サイズは「320px:180px」、「480px:270px」、「800px:450px」、「960px:540px」、「1120px:630px」、「1280px:720px」、「1440px:810px」、「1600px:900px」とより幅広く変更できるようにした。担当されていない通訳区間があれば、「次のタスクに追いつくための早送り」と表示し、次のタスクまで残り何秒か表示させた。担当されていない通訳区間がなければ「遅延を取り戻すための早送り」とし、ライブの再生位置まで残り何秒か表示させた。これら以外の変更点はない。



図 5.1 通訳画面

5.2. 実験目的

手話文字通訳や連係入力、動画操作に慣れるのにどの程度の訓練を要するか明らかにするために、訓練用動画の手話文字通訳を行ってもらい、慣れの主観評価を行う。白石らの研究[10]と本システムの比較を行うために、実験用動画の手話文字通訳を行ってもらい、通訳品質の主観評価や通訳時間の客観評価、必要人数の推定を行う。予備実験では、遅延を取り戻すための再生速度1.5倍速について全員がちょうど良いと感じていた。しかし、複数の講演者に対して有効であるかは明らかではない。本実験では、複数の講演者に対して複数のワーカが再生速度 1.5 倍速は早く感じるかについて主観評価を行う。ボランティアが継続利用のために、予備実験では通訳の楽しさの主観評価を行った。ところが、システムの使いやすさや理解のしやすさも重要であるが、評価を行っていなかった。本実験では、通訳の楽しさのみならず、SUSの主観評価も行う。

5.3. 実験方法

実験の流れは、まず、実験に参加するワーカに日本語入力スピード測定[18]を用いてタイピング速度を記録する。次に、本システムについて説明を行う。次に、本システムに慣れてもらうために訓練用動画の手話文字通訳を4回行ってもらい、毎回訓練の慣れの主観評価を行なった。訓練用動画は2つあり、訓練1回目と3回目には1つ目の動画、2回目と4回目には2つ目の動画を見てもらった。次に、実験用動画の手話文字通訳を行ってもらい、ログデータを収集する。最後に、SUSとアンケートを回答してもらった。管理画面や通訳画面の表示と操作には Google Chrome ウェブブラウザ[19]を利用した。手話動画は事前に保存した動画を YouTube にアップロードして利用した。実験を実施した日程は2020年12月3日、4日、10日、11日、17日の5日間である。本システムの説明方法について、通訳画面にある全ての機能の説明文や通訳区間の間の重複、連係入力、実験開始方法と終了方法を記述した PDF ファイルを読んでもらった。また、ワーカには分からない手話があれば無視して良いことや、リラックスして入力して良いと伝えた。しかし、3回目の実験の前の訓練3回目の後、ワーカ全員が参加しているグループチャットで、あるワーカから「次はあなたの番です。」と表示されているときに、入力してそのままエンターキーを押している人がいると思います。」と意見があった。つまり、担当する通訳区間の一つ前の通訳区間の手話文字通訳を行なっているワーカがいたことである。意見を聞いた後、「通訳中の指示が与えられてから通訳を始めてください」と伝え、3回目実験以降1つ前の通訳区間の手話文字通訳を行うワーカはいなくなった。

5.4. 実験参加者

実験参加者であるワーカは、ろう者・難聴者の大学生 10 名募集し、可能な限り全ての実験に参加してもらうようにした。募集条件は PC を持ち、手話を読めることであり、タイピング速度は問わないとした。これらの条件はいずれも参加者の自己申告とした。本システムはクラウドソーシングを用いるため、ワーカは自由参加を想定している。しかし、手話を読めないワーカが参加した場合、通訳品質は低下するが、その対処方法は今後の課題とした。10 名募集した理由は、予備実験での

必要人数は 15 名であったが、募集が困難のため、ワーカの通訳時間の平均を基準に式 1 で計算すると $Nw = 7.7(Tt = 25.9, Ts = 8s(1\text{秒重複のため}), Rc = 1.5\text{倍速})$ であり、できるだけ 8 名以上となるように募集したからである。その結果、表 5.1 の通り、各実験の参加者人数は最小で 7 名、最大で 9 名であった。また、実験時、訓練前にタイピング速度を測定した結果、全ての実験の平均は 2.2CPS(SD 0.5CPS)であった。実験後、アンケートで手話歴を聞いた結果、10 名のワーカの平均は 10.9 年(SD 7.1 年)、最小で 1 年、最大で 20 年であった。

表 5.1 各実験の実験参加者数とタイピング速度の平均と標準偏差

	ワーカ人数(名)	タイピング速度の平均(CPS)	タイピング速度の標準偏差(CPS)
1回目	7	2.2	0.6
2回目	8	2.3	0.5
3回目	9	2.2	0.5
4回目	9	2.3	0.6
5回目	8	2.3	0.6
平均	8	2.2	0.5

5.5. 手話動画と通訳区間の長さ

実験用動画の長さは、通訳区間 100 区間となるように 7 分 30 秒とした。5 人の講演者に 7 分 30 秒超えるように動画撮影してもらい、5 人分の動画を用意した。講演者には動画撮影の際に、講演しながら原稿などの資料等を見ても構わないが、その資料等が画面に映らないように依頼した。理由は、原稿または発表スライドが動画に映ると、表示されている文字をそのまま写して入力するワーカが現れると想定できるからである。実験用動画の内容は、講演者自身の卒業研究の内容とした。講演者らは情報科学関係の分野を学んでおり、つまり実験用動画は同分野に関する内容である。講演のテーマは以下の通りである。

- ・実験1回目: 発話・発音訓練支援システム
- ・実験2回目: 算数・数学における手話表現
- ・実験3回目: スマートグラスを用いた情報揭示方法
- ・実験4回目: 映像における字幕表示方法
- ・実験5回目: 情報保障の字幕に関する研究

実験参加者の中には、卒業研究の発表を聞いた者もいるが、現実的にワーカ全員が発表に関する知識を持たないとは限らないため、内容を知っている者がいても良いこととした。

講演者には、日本語対応手話を意識するように伝えた。しかし、撮影してもらった動画を確認すると、2回目の講演者以外は中間手話であった。日本語対応手話の場合、助詞は指文字で表し、中間手話では指文字を省くことが多い。2回目の講演者のみ助詞を指文字で表すことが頻繁であ

り、他の講演者は指文字を省くことが多かった。5回目の講演者以外は口形を用いて話していた。つまり、ワーカは手話を日本語に通訳せず、口形を読み取って入力することが可能であった。5回目の講演者のみ口形をあまり使用していなかったため、ワーカは手話を日本語に通訳する必要があった。今回は、日本手話は明らかに日本語と文法の異なるため、今後の課題とした。

各講演者の手話の速さの評価を行った結果を表 5.2 に示す。評価方法は予備実験の時と同様である。結果、1 秒間の手話の速さの平均値は 1.3 単語、最大値は 1.5 単語、最小値は 1.2 単語であった。通訳区間 6 秒の時の手話の速さの平均値は 8.0 単語、最大値は 9.0 単語、最小値は 7.3 単語であった。NHK 週間手話ニュースの映像[20]では、手話の速さは 1 秒で約 1.6 単語であり、それより遅かった。

表 5.2 各講演者の手話の速さ

	手話の速さ (秒/単語)	6 秒の時の手話 の速さ(単語)
1回目	1.5	9.0
2回目	1.4	8.5
3回目	1.3	7.7
4回目	1.2	7.4
5回目	1.2	7.3
平均	1.3	8.0

訓練用動画は講演者の内容とした。1 人の講演者に 2 つの動画を用意してもらった。訓練用動画の長さは、1 つ目は約 7 分、2 つ目は約 6 分 10 秒であった。内容は講演者自身が考え、発表してもらった。講演者に、手話の種類は日本語対応手話を意識するように伝えた。しかし、動画を確認すると中間手話であった。理由は、助詞に指文字を使うことが少なかったからである。

通訳区間の長さは 6 秒、通訳区間の間の重複時間は 1.5 秒とした。6 秒の理由は、4章の予備実験では、9 秒では時間がかかる者がいたからである。そこで、同研究室の学生と指導教員と合わせて 6 名で 9 秒の長さで手話文字通訳を行ってもらった。すると、長く感じるとの意見があった。9 秒から 6 秒に減らして、手話文字通訳を行ってもらうと、短く感じたと肯定的な意見があった。そのため、6 秒と決めた。6 秒以下としなかった理由は、通訳区間を重複しているため、通訳区間が短いほど重複の範囲が大きくなり、入力しなくて良い場面が増えるからである。重複時間 1.5 秒の理由も、6 名に重複時間 1 秒で試してもらうと、後のワーカの入力始めが遅く感じると意見があったからである。通訳区間の長さを 6 秒、重複時間 2 秒で試してもらうと、前後のワーカと被って入力できない場合があったと意見があったからである。タスク完了後の遅延を取り戻すための再生速度は、1.5 倍速とした。理由は4章の予備実験でワーカ全員が 1.5 倍速でちょうど良いと感じたからである。

5.6. システムユーザビリティスケール(SUS) [22]

ウェブサイトは通訳画面のことを指し、SUS の質問項目は以下の通りである。

- (1) このウェブサイトをしばしば利用したいと思う。
- (2) このウェブサイトを利用するには説明が必要となるほど複雑であると感じた。
- (3) このウェブサイトは容易に使いこなす事ができると思った。
- (4) このウェブサイトを利用するのに専門家のサポートが必要だと感じる。
- (5) このウェブサイトにあるコンテンツ(動画、指示文、入力枠、動画操作機能など)は十分に統一感があると感じた。
- (6) このウェブサイトでは一貫性のないところが多々あったと感じた。
- (7) たいていの人は、このウェブサイトの利用方法をすぐに理解すると思う。
- (8) このウェブサイトはとても操作しづらいと感じた。
- (9) このウェブサイトを利用できる自信がある。
- (10) このウェブサイトを利用し始める前に知っておくべきことが多くあると思う。

これらの質問に対し、5 段階評価で「1=同意しない」、「5=同意する」で回答してもらう。項目番号が奇数であれば、数値が大きいほど肯定的である。偶数であれば、数値が大きいほど否定的である。

5.7. アンケート

5.7.1. 手話文字通訳システムの慣れ

手話文字通訳システムの訓練の慣れのアンケート内容は以下の通りである。

- ・「手話文字通訳(手話を読み取りながら手話を日本語に通訳しながら文字を入力する作業)の慣れはどうか？」
- ・「ボタンまたはショートカットキーによる動画操作(1 秒戻る、再生速度調整、再生停止)の慣れはどうか？」
- ・「前後の入力文に欠落や重複があるか確認することの慣れはどうか？」

これらの質問に対し、5 段階評価で「1=慣れない」、「5=慣れた」とした。

5.7.2. 前後入力文を確認したか

アンケートで「通訳中に前後の入力文から欠落や重複があるか確認しましたか？(以下、前後の入力文を確認したか)」を 3 段階評価で「1=確認しなかった」、「2=確認しない時があった」、「3=毎回確認した」と回答してもらった。

5.7.3. 再生倍率 1.5 倍速で遅延を取り戻す方法

アンケートで「通訳後、遅れを取り戻すために再生倍率を 1.5 倍速にしましたが、速く感じましたか？(以下、再生倍率 1.5 倍速は速く感じたか)」を 5 段階評価で「1=速い」、「3=ちょうど良い」、「5=遅い」と回答してもらった。

5.7.4. 通訳の楽しさ

アンケートで「通訳は楽しかったですか(以下、通訳の楽しさ)」を 5 段階評価で「1=楽しくなかった」, 「5=楽しかった」と回答してもらった。

5.8. 通訳品質の評価方法

通訳品質は実験に参加していない部外者 7 名に分析を依頼した。各実験の各通訳区間の通訳結果を実験参加者以外の 3 名に評価してもらった。同評価者は講演者と同様に情報科学関係の分野を学んでおり手話が読み取れる者とした。これは、講演内容が情報科学関係の卒業研究のため、その分野の知識を要するからである。評価者らの手話歴は、最小 1 年 8 ヶ月、最大 20 年であった。通訳品質の評価者は、本来は長年手話でコミュニケーションしてきた者を採用した方がよかったが、今回は自己申告で募集したため、2 年未満の者が含まれてしまった。今後は、ワーカの属性を考慮し、評価者の募集条件の見直しを行う。

評価の流れは、まず、評価方法の説明を行った。次に、講演の内容理解のために手話動画を 1 回見てもらった。最後に、通訳品質の主観評価を行ってもらった。通訳品質の評価方法は、通訳区間の手話動画と担当したワーカの入力結果を見比べ、白石らの研究[10]を参考に、以下の 7 段階評価で行なった。

- (1) 意味伝達率ほぼ 0%
- (2) 意味伝達率 20%未満
- (3) 意味伝達率 20%以上 40%未満
- (4) 意味伝達率 40%以上 60%未満
- (5) 意味伝達率 60%以上 80%未満
- (6) 意味伝達率 80%以上
- (7) 意味伝達率ほぼ 100%

例えば、評価者が意味伝達率 20%未満だと感じた場合、通訳品質は 2 となる。白石らの研究と本研究での通訳品質の評価方法の違いは、前者は通訳区間を意味単位で区切るため、通訳区間 1 区間に対して評価を行うことができる。しかし、後者では通訳区間を重複するため、現在のワーカの入力結果のみならず、前後のワーカの入力結果の繋がりを確認した上で、点数を付けてもらった。繋がりについて、具体的には、欠落があれば評価を低くしてもらい、重複があったとしても正しい文であれば評価を低くしなくても良いと伝えた。何も入力してなかった時も同様であり、前後のワーカが自分の担当区間の文を入力していた場合、評価は下げなくて良いと伝えた。また、欠落や重複以外にも、現在のワーカの入力結果を見て、誤りがあった時や、助詞だけ落ちていた時は低くしても良いと伝えた。評価者が手話の見落としした時を想定して巻き戻しを使用できるようにした。

5.9. 実験結果

5.9.1. 手話文字通訳システムの慣れ

実験の前に訓練を4回行い、毎回慣れについて5.7.1節のアンケートを回答してもらった。実験は5回行ったため、例えば全ての実験に参加した者は訓練4回と実験回数5回を掛けて最大20回行ったことになり、3回参加した場合は12回となる。図5.2と図5.3、図5.4に手話文字通訳と連係入力、動画操作の慣れについての結果を示す。棒グラフが平均値でエラーバーが標準偏差を示している。手話文字通訳や連係入力、動画操作の慣れのフリードマン検定を行った結果、有意差があったため、ホルム法を用いて多重比較を行った結果を同図に示した。ワーカの慣れの平均を折れ線グラフで、標準偏差をエラーバーで示した。ワーカの人数は訓練1回目から8回目の間では10名、9回目から16回目では8名、17回目から20回目では6名である。実験参加者の訓練回数に着目した結果であり、ある訓練から次の訓練の期間については考慮していないが、訓練4回の間隔で、1日や6日、7日後に次の実験に参加したワーカもいる。

手話文字通訳の慣れにおいて、訓練回数1回と2回以降全て有意差($p<.01$)、2回と4回は有意差($p<.05$)、2回と6回以降全て有意差($p<.01$)であった。全員「慣れた」と回答したのは訓練12回であり、訓練3回から12回の間には有意差は認められなかった。訓練1回では慣れない傾向であり、3回以上行くと慣れる傾向であることが分かった。訓練4回の後、実験を行い、別の日に5回から8回までの訓練を行っていたため、訓練2回と5回では差はなく、2回と6回では差があった。最初の訓練で「1=慣れない」と答えた者のコメントは以下の通りである。

- ・「手話の速さと同時にタイピングが難しいから」
- ・「第一言語が音声日本語で、日本語対応手話(中間手話)から日本語に通訳するのに時間を要する。」

訓練4回後、「5=慣れた」と答えた者からコメントは以下の通りである。

- ・「手話に集中してもキーボードで入力できるようになった。」
- ・「いままで私のみ入力する所を見たが、4回目に、上と下の文章に合わせて入力した。」

連係入力の慣れにおいて、訓練回数1回と2回のみ有意差($p<.05$)であり、1回目と3回以降全て有意差($p<.01$)であった。全員「慣れた」と回答したのは訓練14回であり、訓練2回から14回の間には有意差は認められなかった。訓練1回では慣れない傾向であり、2回以上行くと慣れる傾向であることが分かった。

最初の訓練で「2=やや慣れない」と答えた者からコメントは以下の通りである。

- ・「とにかくスピード優先でしてしまった」

訓練4回後で「4=やや慣れた」と答えた者からコメントは以下の通りである。

- ・「前の文を見てここは消そうと余裕を持つことができた。」
- ・「入力している時に変換ミスがないか確認すると、手話を見ることと重複がないかの確認ができないところがあるから」

「5=慣れた」と答えた者のコメントは以下の通りである。

・「4回もしたので慣れました。2 回無入力をしてしまいました…。(上と下の入力者に挟まれて入力する文章がなくなったので)」

・「数を重ねたので慣れたと思う。」

連係入力の慣れにおいて、訓練回数 1 回と 2 回のみ有意差 ($p<.05$) であり、1 回目と 3 回以降全て有意差 ($p<.01$) であった。全員「慣れた」と回答したのは訓練 14 回であり、訓練 2 回から 14 回の間には有意差は認められなかった。訓練 1 回では慣れない傾向であり、2 回以上行くと慣れる傾向であることが分かった。

動画操作の慣れにおいて、訓練回数 1 回と 6 回以降では有意差 ($p<.01$) であった。全員慣れたと回答したのは訓練 18 回であり、訓練 6 回との間に有意差は認められなかった。訓練 6 回行くと慣れる傾向であることがわかった。最初の訓練で「1=慣れない」と答えた者のコメントは以下の通りである。

・「動画操作は使わなかったため。」

・「停止などはあまり使わなかったから…。動画の速度はタスク待機中に変えられるので。」

訓練 4 回後で「5=慣れた」と答えた者のコメントは以下の通りである。

・「速度調整を一度したらもう使わないため。」

また、「1=慣れない」と答えた者のコメントは以下の通りである。

・「使い道がまだわからないです、ごめんなさい。遅延しているので、今の時間にスキップしたいのですが、そのボタンがないのでほしいと思いました。」

・「使わなかったので…」

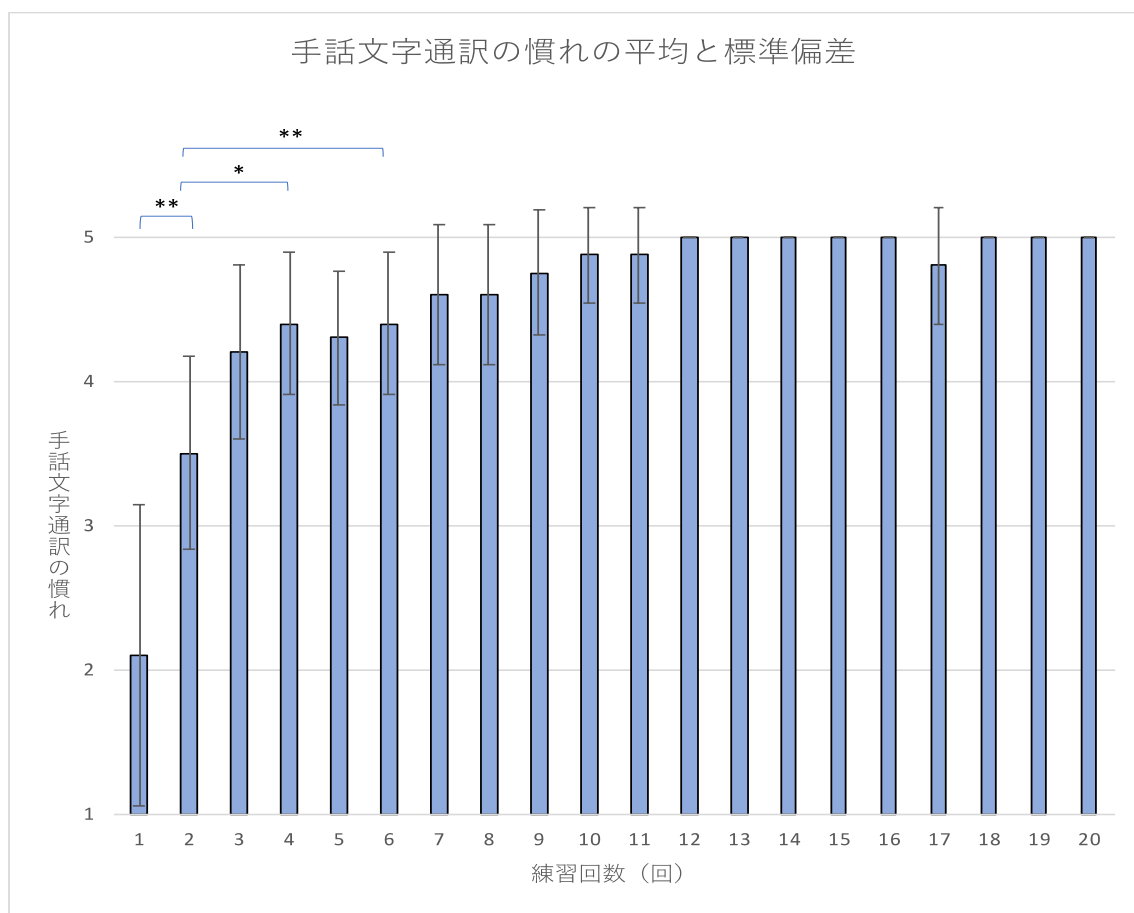


図 5.2 手話文字通訳の慣れの平均と標準偏差

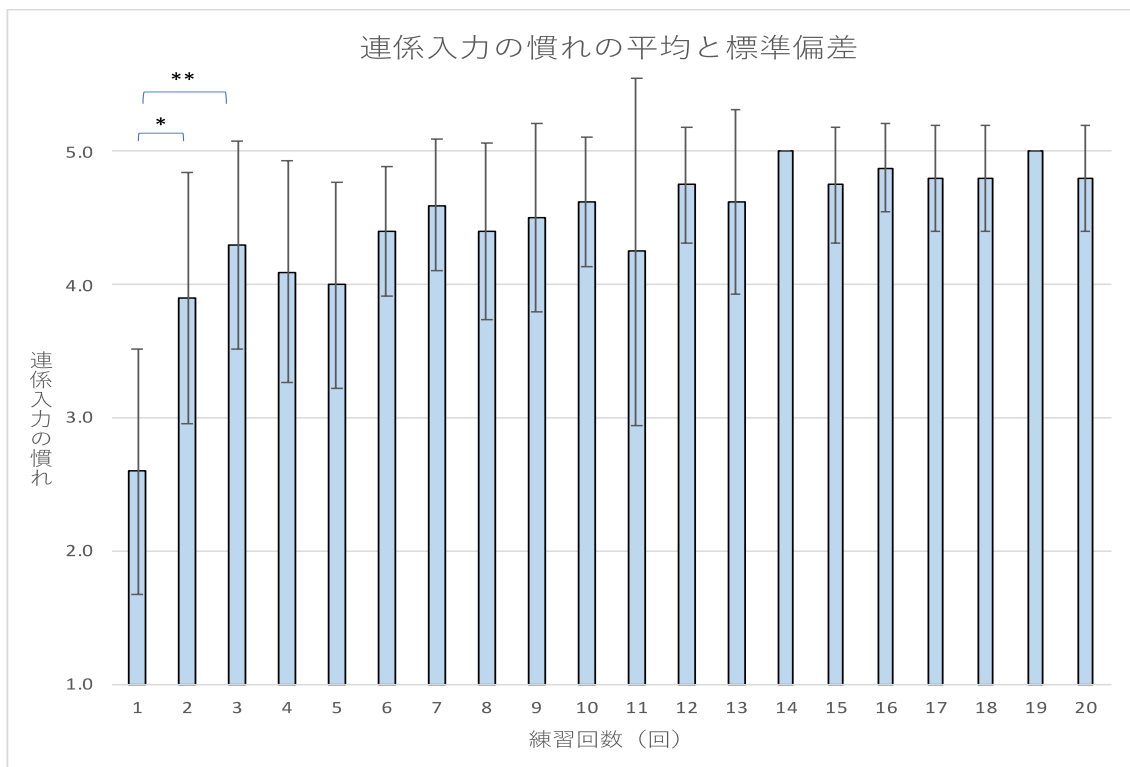


図 5.3 連係入力の手順の慣れの平均と標準偏差

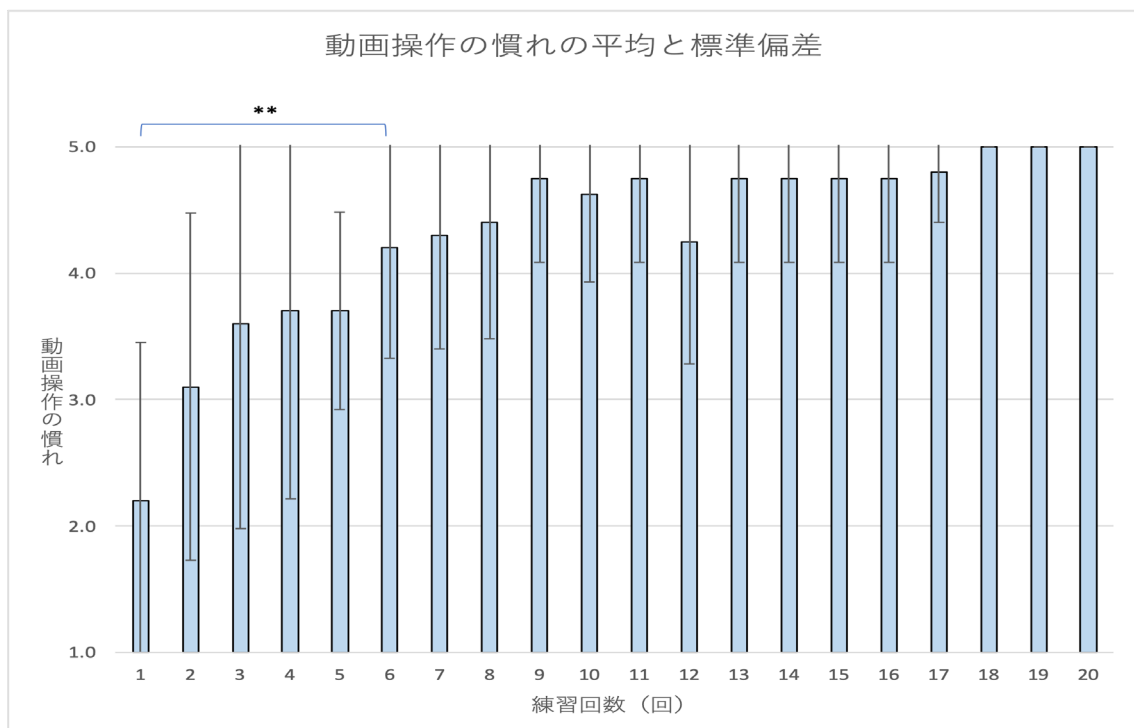


図 5.4 動画操作機能の手順の慣れの平均と標準偏差

5.9.2. 通訳時間と必要人数

各実験での通訳時間の平均と標準偏差, ワーカ人数, 必要人数を表 5.3 に示す. 通訳時間の計算の際, 1回目で1区間, 2回目で1区間, 4回目で1名のワーカが担当した全ての通訳区間3区間を外れ値として除外した. 1回目と2回目では, システムの不具合によりタスクを完了できなかったからである. 4回目の1名のワーカを除外した理由は, 通訳3回目でシステムの不具合により, 途中で抜けて以降, タスクが割り当てられることはなく, 通訳回数が2回と少ないため, 失敗と見なしたからである. 結果, 通訳時間の全体の平均は17.0秒(SD 8.0秒)であり, 5回目が最大で19.0秒(SD 8.2秒), 2回目が最小で14.3秒(SD 7.1秒)であった. ワーカ人数は1回目実験では7名, 2回目と5回目実験では8名, 3回目と4回目実験では9名であった. 必要人数は各実験の通訳時間の平均を基準に, 通訳区間の時間は6秒, 重複時間は1.5秒, 早送りの再生速度は1.5倍速とし, 計算後にその数値を切り上げた. その結果, 最大で11名, 2回目実験が最小で8人であった.

全ての実験での各ワーカの通訳時間の平均を図 5.5 の通りにヒストグラムで表示させた. 縦軸はワーカ人数, 横軸は通訳時間の階級であり, 例えば0~4は0秒以上5秒未満, 5~9は5秒以上10秒未満とした. 結果, 15秒以上20秒未満が最大で13名であった. 次点で20秒以上25秒未満では11名であった. 25秒以上30秒未満は6名, 5秒以上10秒未満は2名, 35秒以上40秒未満は1名であった. 通訳区間の長さは6秒であるが, ログを分析すると6秒未満で完了することがあったワーカは2名中1名いた. そのワーカはルールの不理解により, 担当する通訳区間の一つ前の通訳区間の手話文字通訳を行っていたため, 平均値が低かった.

表 5.3 通訳時間とワーカ人数と必要人数

	通訳時間の平均(秒)	通訳時間の標準偏差(秒)	ワーカ人数(名)	必要人数(名)
1回目	16.6	7.4	7	10
2回目	14.3	7.1	8	8
3回目	17.0	8.9	9	10
4回目	18.3	8.3	8	11
5回目	19.0	8.2	9	11
平均	17.0	8.2	8	10

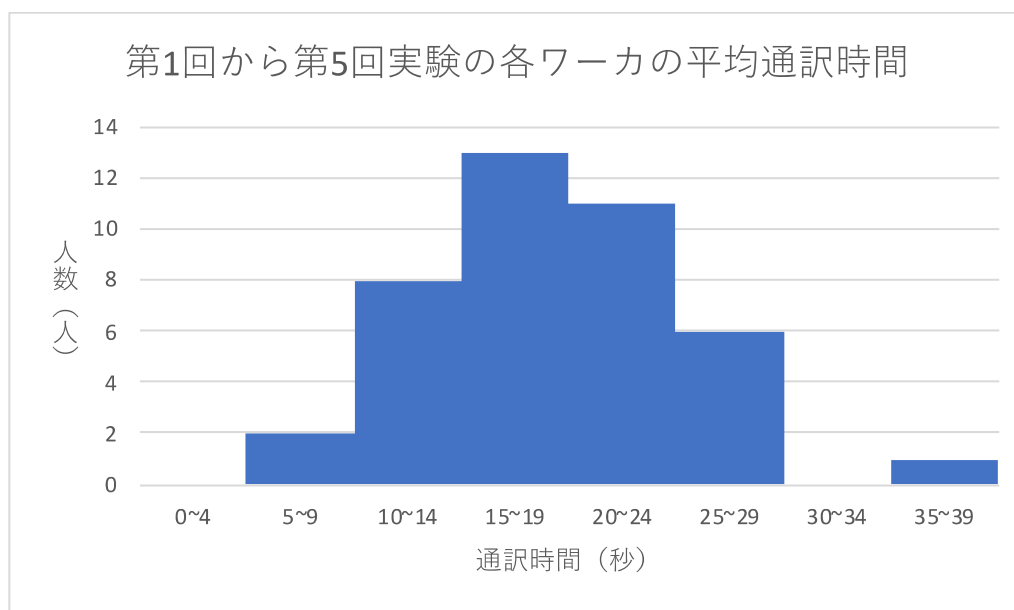


図 5.5 各ワーカーの平均通訳時間のヒストグラム

5.9.3. 通訳回数

各実験でのワーカー人数と通訳回数の平均と標準偏差を表 5.4 に示す。通訳回数の計算の際、5.9.2 節と同様の理由で、1回目で 1 区間、2 回目で 1 区間、4回目で 1 名のワーカーが担当した全ての通訳区間 3 区間を外れ値として除外した。結果、1回目ではワーカー人数が 7 名と一番少なく、通訳区間数が 99 区間の時、通訳回数の平均は 14.1 回 (SD 5.5 回) であった。2 回目ではワーカー人数は 8 名、通訳区間数が 99 区間の時、通訳回数の平均は 12.4 回 (SD 6.2 回) であり、標準偏差が一番高かった。3 回目と 5 回目はワーカー人数が 8 名、通訳区間数は 100 区間、どれも同数であるため、平均値は 11.1 回と同じであった。4 回目ではワーカー人数が 8 名、通訳区間数が 97 区間の時、平均値は 12.1 回 (SD 4.1 回) であった。通訳回数が最大のワーカーは、2 回目で最大で 27 回であった。通訳回数が最小のワーカーは 1 回目と 2 回目で 4 回であったが、2 名のワーカーはシステムの不具合により途中で抜けたワーカーである。途中で抜けなかったワーカーで最小のワーカーは 3 回目で 5 回であった。

表 5.4 1回目から 5 回目の通訳回数の平均と標準偏差

	ワーカー人数(人)	通訳区間数(区間)	通訳回数の平均(回)	通訳回数の標準偏差(回)	通訳回数の最大値(回)	通訳回数の最小値(回)
1回目	7	99	14.1	5.6	25	4
2回目	8	99	12.5	6.2	27	4
3回目	9	100	11.1	4.4	20	5
4回目	8	97	12.1	4.1	21	8
5回目	9	100	11.1	3.8	20	6

5.9.4. 通訳品質と通訳区間の間の文の欠落と重複のある通訳区間の割合

各実験の通訳品質の平均と標準偏差, ケンドールの一致係数, 通訳区間の間の文の欠落率と重複率を表 5.5 に示す. 通訳品質の計算の時, 5.9.2 節と同様の理由で, 1回目で 1 区間, 2 回目で 1 区間, 4回目で 3 区間を外れ値として除外した. その結果, 通訳品質の平均は 6.3(SD 1.3), 5 回目まで最大で 6.6(SD 0.9), 2 回目が最小で 5.8(SD 1.7)であった. 評価者のばらつきを評価するために, ケンドールの一致係数を用いて一致度の評価を行なった. ケンドールの一致係数の数値の範囲は 0 から 1 であり, 1 に近いほど一致していて, 0 に近いほど一致していない. 結果, 1 回目は 0.545 ($p<.01$), 2 回目は 0.65 ($p<.01$)であった. 3 回目では 0.263, 4 回目では 0.217, 5 回目では 0.205 であり, 有意差は認められなかった.

通訳区間の間に文の欠落や重複があった時にカウントし, 欠落率と重複率を求めた. 調査方法は 4.3.2 節と同様である. 欠落率や重複率の計算の時, 前述した外れ値と最後の通訳区間を省いた. 後者の理由は, 通訳区間は 100 区間の場合, 最後の通訳区間は次の通訳区間がないため, 省くことで 99 区間となる. 結果, 欠落率の平均は 19%(SD 7%), 3 回目まで最大で 26%, 4 回目が最小で 9%であった. 重複率は平均で 6%(SD 2%), 2 回目まで最大で 10%, 3 回目が最小で 4%であった.

表 5.5 通訳品質と通訳区間の間の文の欠落率と重複率

	通訳品質 の平均	標準偏差	ケンドール の一致係数	欠落率	重複率
1回目	6.0	1.7	0.545	23%	7%
2回目	5.8	1.7	0.65	25%	9%
3回目	6.4	1.1	0.263	26%	4%
4回目	6.6	0.7	0.217	9%	4%
5回目	6.6	0.9	0.205	12%	5%
平均	6.3	1.3		19%	6%

5.9.5. 前後の入力文を確認したかの主観評価結果

5.7.2 節の前後の入力文を確認したかについて回答してもらった結果の平均と標準偏差を表 5.6 に示す. 5.9.2 節と同様の理由で, 4 回目に参加した 1 名のワーカのみ外れ値として除外した. 結果, 全ての実験において全体の平均値は 2.8(SD 0.5)であった. つまり, ほとんどのワーカが確認していることがわかった. 「確認しなかった」と答えたワーカは 2 回目で 1 名のみであり, 「確認しない時があった」と答えたワーカは各実験で 1 名から 2 名であった.

「確認しなかった」と答えた者のコメントは以下の通りである.

・「ついていくのに精一杯だったから」

「確認しない時があった」と答えた者のコメントは以下の通りである.

- ・「気を付けていますが、実験中に 1～2 回は確認忘れてしまいます…。少し慌ててしまいます」
- ・「タイピングが遅かったときは確認しない時があった」

「毎回確認した」と答えた者のコメントは以下の通りである。

- ・「できるだけ情報を正確に提供したかったから」
- ・「文脈は大丈夫かどうかの確認をしたから」
- ・「余裕があったから」

表 5.6 前後入力文を確認したかの主観評価結果

	前後入力文を確認したかの平均	標準偏差
1 回目	2.7	0.5
2 回目	2.6	0.7
3 回目	2.8	0.4
4 回目	2.9	0.3
5 回目	2.9	0.3
平均	2.8	0.5

5.9.6. 再生倍率 1.5 倍速は速く感じたかの主観評価結果

5.7.3 節の再生倍率 1.5 倍速は速く感じたかを回答してもらった結果の平均と標準偏差を表 5.7 に示す。5.9.2 節と同様の理由で、4 回目に参加した 1 名のワーカのみ外れ値として除外した。結果、全ての実験において全体の平均値は 3.0(SD 0.6)であった。つまり、ほとんどのワーカがちょうどいいと回答していた。やや遅く感じた、遅く感じたと回答した者は全ての実験において 1 名のみであった。やや速く感じたと回答した者は 3 回目以外で 1 名いた。速く感じたと回答した者は 2 回目のみ 1 名いた。

「やや遅い」と答えた者のコメントは以下の通りである。

- ・「非常に遅れた時でも 1.5 倍速だからじれったかった。」
- ・「今回の本番のときは、話している人のスピードが遅かったのと、間が多く合ったので、人に合わせて手動で変えるほうがいいかもしれません」

「やや速く感じた」と答えた者のコメントは以下の通りである。

- ・「少し早くて全部の手話はわからなかった。所々で読み取った。」

「速く感じた」と答えた者のコメントは以下の通りである。

- ・「手話が読み取れなかったから」

「ちょうど良い」と答えた者のコメントは以下の通りである。

- ・「早送りではあるが、早すぎなかった。」
- ・「理解できるスピードだったから」

表 5.7 再生速度 1.5 は速く感じたかの平均と標準偏差

	再生速度 1.5 倍速は速く感じたかの平均	標準偏差
1回目	3.0	0.6
2回目	2.8	0.6
3回目	3.1	0.5
4回目	3.1	0.8
5回目	3.0	0.9
平均	3.0	0.6

5.9.7. 動画操作機能の主観評価結果

「この中で使った機能はどれですか(複数可)(以下、動画操作に使用した機能)」を以下の 4 つの項目を複数回答で回答してもらった。

- ・「再生倍率調整」
- ・「1 秒戻る」
- ・「再生・停止」
- ・「特になし(再生倍率調整はデフォルトで 0.8 倍速)」

結果を図 5.6 に示す。回答結果を横棒グラフで示した。動画操作機能の中から一番使われた機能が「再生倍率調整」であり、全ての実験で6名であった。「再生・停止」を使用した者は5回目で1名のみ、「1 秒戻る」は 0 名、「特になし」は各実験で 1 名から 3 名であった。「再生倍率調整」を使用した者からのコメントは以下の通りである。

- ・「担当部分の再生がループされるが、遅いと待たなくてはいけなくなるため、×1.0 にしていた。」
- ・「0.8 倍だと微妙に追いつけない。0.6 倍なら一回の再生でほぼ入力が間に合うから」
- ・「0.8 倍だと見逃しがたまにあったので 0.5～0.7 くらいにしていた」

「特になし」と答えた者から「ちょうど良かった。あまり早すぎてもわからないし、遅すぎても打ちにくいから」などの回答があった。「再生・停止」と答えた者から「一旦停止して手指の位置を確認、入力という作業が必要だったため。」と回答があった。

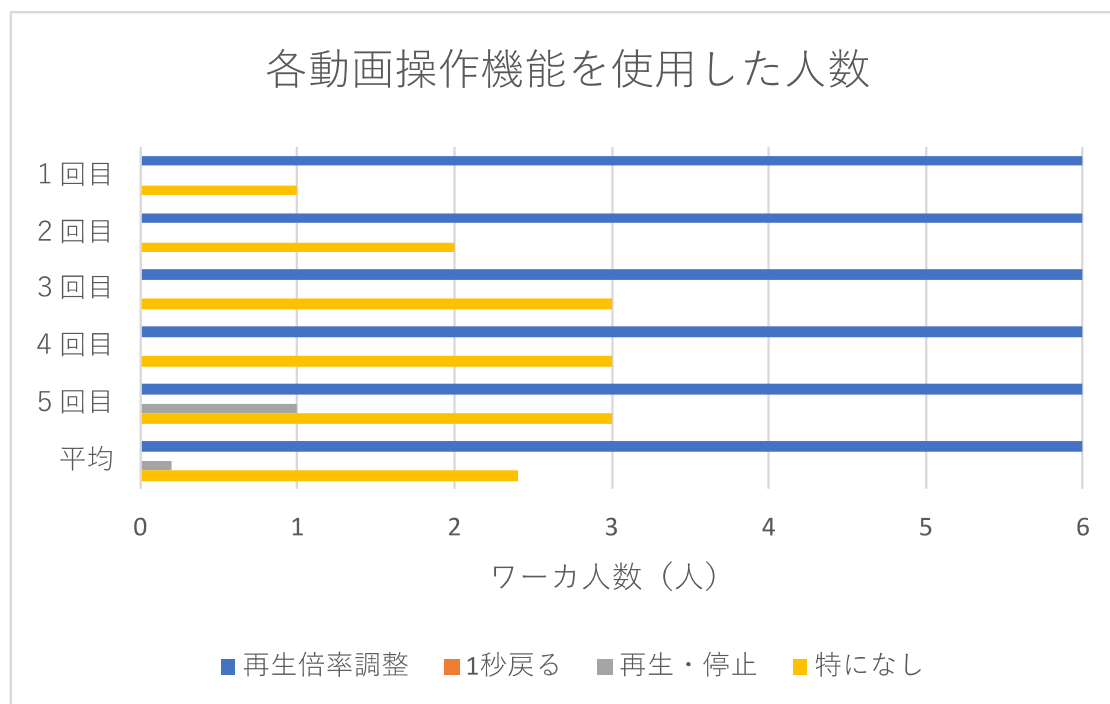


図 5.6 「この中で使った機能はどれですか？」の主観評価結果

5.9.8. 通訳の楽しさの主観評価結果

5.7.4 節の通訳の楽しさを回答してもらった結果の平均と標準偏差を表 5.8 に示す。5.9.2 節で述べた同様の理由で、4 回目に参加した 1 名のワーカーのみ外れ値として除外した。結果、全ての実験において通訳の楽しさの平均は 4.2 以上であった。つまり、ほとんどのワーカーが通訳を楽しめたことがわかった。やや楽しくなかったと答えた者は 2 回目で 1 名のみであった。どちらでもないと回答した者は 1 回目と 3 回目、4 回目で 1 名、5 回目で 2 名であった。

「楽しかった」と答えた者のコメントは以下の通りである。

- ・「助けられる側ではなく、助ける側に回れたことが心嬉しく感じる。」
- ・「通訳は好きであるため」

「どちらでもない」と答えた者のコメントは以下の通りである。

- ・「今回の手話が読み取りにくかったから」

「やや楽しくなかった」と答えた者のコメントは以下の通りである。

- ・「とても疲れたから」

表 5.8 通訳の楽しさの主観評価結果

	通訳の楽しさの平均	標準偏差
1 回目	4.4	0.7
2 回目	4.3	1.0
3 回目	4.2	0.6
4 回目	4.5	0.7
5 回目	4.4	0.8
平均	4.4	0.8

5.9.9. SUS による評価

各実験後、各ワークに SUS を用いて主観評価を行ってもらい、回答結果から点数への計算を行い、各実験での点数の平均と標準偏差を表 5.9 に示す。点数の見方はウェブサイト[23]を参考にすると、SUS の標準平均は 68 点であり、68 点未満から 51 点以上の場合、ユーザビリティは「やや悪い」と評価することになる。全体の平均値は 62.8 点 (SD 14.2 点) であり、1 回目が最低で 56.4 点 (SD 13.6 点)、5 回目が最大で 66.4 点 (SD 14.5 点) となり、「やや悪い」結果であった。

5 段階評価で「1=同意しない」、「5=同意する」で回答してもらった結果を表 5.10 に示す。奇数の項目番号で平均的に 4 以上 5 以下、偶数の項目番号で平均的に 1 以上 2 未満と答えた項目番号は(1),(5),(6),(9)であり、これらは肯定的な傾向であった。奇数の項目番号で平均的に 1 以上 2 未満、偶数の項目番号で平均的に 4 以上 5 以下と答えた項目番号は(10)であり、これは否定的な傾向であった。残り(2),(3),(4),(7),(8)はどちらでもない傾向であった。

表 5.9 SUS の評価結果

	SUS の平均(点)	SUS の標準偏差(点)
1 回目	56.4	13.6
2 回目	64.4	13.0
3 回目	61.9	13.2
4 回目	63.8	14.8
5 回目	66.4	14.5
平均	62.8	14.2

表 5.10 SUS の質問項目ごとの平均点と標準偏差

項目番号	平均	標準偏差
(1)このウェブサイトをしばしば利用したいと思う.	4.3	0.6
(2)このウェブサイトを利用するには説明が必要となるほど複雑であると感じた.	3.4	1.3
(3)このウェブサイトは容易に使いこなす事ができると思った.	3.7	1.0
(4)このウェブサイトを利用するのに専門家のサポートが必要だと感じる.	3.3	1.2
(5)このウェブサイトにあるコンテンツ(動画、指示文、入力枠、動画操作機能など)は十分に統一感があると感じた.	4.2	0.9
(6)このウェブサイトでは一貫性のないところが多々あったと感じた.	1.8	0.8
(7)たいていの人、このウェブサイトの利用方法をすぐに理解すると思う.	3.8	0.8
(8)このウェブサイトはとても操作しづらいと感じた.	2.1	1.0
(9)このウェブサイトを利用できる自信がある.	4.0	0.9
(10)このウェブサイトを利用し始める前に知っておくべきことが多くあると思う.	4.2	0.9

5. 10. 考察

5. 10. 1. 手話文字通訳の慣れ

4章の予備実験では、連係入力を行った者が 1 名のみであり、他のワーカのコメントを読むと、手話文字通訳に精一杯であることが分かった。訓練量が足りなかったことが原因だと考えられる。本実験では、訓練用動画の手話文字通訳を4回行ってもらい、毎回慣れの主観評価を行なった。結果、手話文字通訳の慣れにおいて、最初の訓練後では慣れない傾向であることが分かった。慣れない者のコメントを見ると、手話を見ながらタイピングを行うことが難しい、または第一言語が音声日本語のため、手話から日本語への通訳に時間を要することが理由であった。このことから、初めて手話文字通訳を行う者は、手話と文字を交互に見ながらタイピングを行うことや、手話から日本語への通訳に慣れていない者が多いことが分かった。訓練 3 回行った後の慣れは、全員慣れたと回答した訓練 12 回との有意差が認められなかったため、最低 3 回は必要であることが分かった。

訓練を4回行ったワーカのコメントを見ると、手話を見ることに集中しながらタイピングできるようになったこと、訓練3回までは通訳区間の手話文字通訳に集中していたが、訓練4回では連係入力を意識したと回答した者がいた。つまり、これらを踏まえると、手話文字通訳に慣れていない者は、訓練を積み重ねるほど手話文字通訳に余裕が出て連係入力を意識できるようになることが考えられる。よって、今回の実験からは手話文字通訳の慣れに要する訓練回数は最低 3 回であり、訓練時間は約6分10秒の動画を 1 回、約7分の動画を2回、手話文字通訳を行ってもらったため、最低約20 分だと考えられる。

訓練2回と5回では有意差がなく、訓練2回と6回以降では有意差があった。5回以降は他の日に実験を行ったため、慣れが低下していたと考えられる。このことから、久しぶりに手話文字通訳を行うワーカは訓練をした方がいいことが分かった。

5.10.2. 連係入力への慣れ

5.10.1 節で述べた通り、予備実験では連係入力を行なったワーカが1名のみであり、訓練量が足りないことが原因だと考えた。本実験では、訓練用動画を行ってもらい、連係入力への慣れの主観評価を行なった。結果、連係入力への慣れにおいて、最初の訓練後では平均的にやや慣れない傾向であった。コメントを見ると、タスクを早く終わらせることを優先にする者がいることが分かった。つまり、手話文字通訳に集中し、連係入力に意識が向いていなかっただろう。訓練2回から全員慣れたと回答した訓練14回の間では有意差が認められず、訓練回数は最低2回必要であることが分かった。やや慣れたと答えた者のコメントを見ると、前のワーカの文章を見て重複を削除する余裕ができたことと回答していた。慣れたと答えたワーカのコメントでは、数を重ねたため、慣れたと思うと回答していた。つまり、訓練を重ねることでワーカは連係入力を意識できるようになることが分かった。よって、連係入力への慣れに要する時間は約6分10秒の動画を1回、約7分の動画を1回、合わせて最低約13分だと考えられる。

他のコメントを見ると、入力中に変換の誤りがないか確認しながら手話または重複がないか確認することが困難であることが分かった。今後は、修正タスクを設け、文字入力や連係入力を共に行うことに負担を感じるようなワーカのために、修正ワーカに任せる方法を検討する。また、通訳時間を短くし、通訳区間の重複時間を長くしたため、入力箇所が被りやすく、何も入力しなかったと回答したワーカがいた。講演者は常に話し続けているのではなく、時々間を空けて話すため、何も入力しないでいい状況が発生したと考えられる。今後は、動画解析を用いて、手話の合間を見つけて区切る、または話し始める頃にタスクを割り当てる方法を検討する。

5.10.3. 動画操作への慣れ

通訳画面には動画操作機能があり、「再生速度調整」、「再生・停止」、「1秒戻る」を設け、ショートカットキーまたはボタンを用いた動画操作の慣れを評価した。結果、動画操作の慣れにおいて、最初の訓練後では平均的にやや慣れないと答える傾向にあることが分かった。訓練6回以降では、全員慣れたと答えた訓練18回と有意差がなかった。慣れないと答えた者のコメントを見ると、動画操作は使わなかった、停止などはあまり使わなかったことが分かった。このことから、動画操作の使い道が分からなかったから慣れなかったことも1つの原因だと考えられる。訓練4回行なった後、慣れたと答えた者のコメントを見ると、再生速度調整を一度行ったらそれ以降使用しない者がいることが分かった。慣れない者のコメントを見ると、訓練4回終えても動画操作機能の使い道が分からないことが分かった。動画操作機能の主観評価結果を見ると、「再生・停止」は5回目実験で1名が使用しただけであり、「1秒戻る」を使用する者がいなかった。つまり、使い道が分からない機能は「再生・停止」と「1秒戻る」の2つだろう。おそらく、通訳中では、通訳区間を見終えた後、通訳区間

の最初の時間に戻ることを繰り返すため、「再生・停止」と「1 秒戻る」を使わなくても手話を見直せることが原因だと考えられる。つまり、動画操作機能を使わない、またはあまり使わない理由で慣れたか慣れないかの判断が難しかったことが分かった。本実験の説明では、動画操作機能の使用は強制ではないと伝えていなかった。今回は訓練6回以降で慣れる傾向であったが、要らないと思う機能は無視してもいいと説明していれば、より早く慣れることが考えられる。

5.10.4. 通訳時間

予備実験では、通訳区間の時間を 9 秒とした時、通訳時間の平均は約 26 秒(SD 13.1 秒)であり、通訳区間の約 3 倍の時間がかかっていた。受講者は字幕表示の遅れが大きいほど、講演の進行状況の把握が困難となる。本実験では、より字幕の遅れを小さくするために、通訳区間を 6 秒にした。白石らの研究[10]では、日本語対应手話の講演者の時、平均区切り時間 8.8 秒の時、通訳時間は 22.4 秒である。本実験では、2回目の日本語対应手話の講演者の時、通訳時間は 14.3 秒であり、通訳区間の長さは異なるが、字幕の遅れは減少していた。中間手話の講演者の時、平均区切り時間 8.3 秒の時、通訳時間は 19.3 秒である。本実験では、1, 3, 4, 5 回目の中間手話の講演者の時、通訳時間は平均 17.7 秒(SD 8.3 秒)であり、字幕の遅れは減少していた。本研究では、通訳区間の長さは一定時間で前後を重複しているため、平均通訳時間を基にすると、通訳区間の長さに重複時間を引いた時間 4.5 秒後に新しい通訳結果が字幕に表出される。白石らの研究では、意味単位で区切るため、例えば中間手話の場合、平均通訳時間 17.7 秒の平均区切り時間 8.3 秒後に新しい通訳結果が字幕に表出される。また、平均区切り時間がまちまちのため、字幕表出まで時間がかかる場合もある。さらに、1 グループの通訳結果の投票もあるため、最適文が表示されるまでかかる時間を含めると、より時間がかかる。これらの観点を踏まえると、字幕の遅れに関しては、本研究の方が小さいと言える。

各ワーカの通訳時間のヒストグラムによると、全体的に 5 秒から 30 秒未満にタスクを完了するワーカが多かった。1 名はルールの不理解により、平均的に 5 秒から 9 秒未満で終わっていた。理由は、前の通訳区間の手話文字通訳を行い、担当区間で入力を終えていたからである。タスク割り当て前から入力できるようにしていたため、このようなワーカが現れてしまった。今後は、入力開始時間や終了時間の制御を行い、確実にタスク開始時に入力を促すようなインタフェースを検討する。1 名のみ 35 秒以上かかっていた。受講者からして、35 秒は長すぎると考えられる。長くかかった理由を、コメントを読むと、正しく通訳したいと思うワーカもいれば、連係入力を意識して時間がかかるワーカがいることが分かった。本システムでは、通訳タスクのみであったため、正しく通訳しつつ、連係入力を行わないと通訳品質の高い入力結果を字幕に表出できない。これが原因で時間がかかるワーカがいた。今後は、修正タスクを設けて、修正ワーカが通訳結果の誤りを修正する方法によって、通訳ワーカは誤りがあっても修正ワーカに任せられる心理的安全によって通訳時間が短縮できるか検討する。

5.10.5. 必要人数

予備実験では、ワーカ数4名で行なったため、遅延が大きく発生していた。2 連続通訳を担当するワーカがおり、楽しくないと感じていた。本実験では、負担を減らすために、推定された必要人数に近いワーカ数で実験を行なった。しかし、通訳区間の時間 6 秒や重複時間 1.5 秒であり、予備実験の時とは異なるため、必要人数は明らかではない。そこで、実験を行なった結果、必要人数は最小で8名、最大で 11 名であった。白石らの研究[10]では、1 グループ3名以上を 5 グループ以上要し、区切り担当ワーカや投票担当ワーカを含めると、合計で17名以上である。つまり、本システムは、より少ない人数で行うことが可能である。

各実験に参加したワーカの人数と各実験の必要人数を比較すると、2 回目実験以外は必要人数以下であることが分かった。しかし、2 回目実験では 51 区間目でシステムの不具合で一時的に抜けた者がいたため、一時的に必要人数未満となってしまった。また、ログを分析すると、必要人数を満たしていても、通訳区間 15 区間までは遅延なくできていたが、16 区間以降は常に遅延が発生していた。理由は、ワーカの通訳時間にばらつきが大きかったからである。例えば 10 区間までと 100 区間までの通訳時間の平均値は同じではなく、前者と後者の必要人数が異なる場合がある。今後、必要人数の増減を防ぐために、通訳タスクに制限時間を設ける方法を検討する。しかし、制限時間を超えた場合、通訳品質が低下することは想定できるため、修正タスクを設け、制限時間を超えた通訳区間に対して修正できるようにする方法を検討する。

5.10.6. 通訳品質

受講者が字幕を読んで内容を理解するには、通訳品質は高い程良い。白石らの研究[10]では、通訳品質を保障するために、1 つの通訳区間に対し、グループで行う。しかし、人的コストが大きくなるため、本研究では、1 つの通訳区間に対し、1名のワーカで行う。さらに、手話の見落としや文脈の理解のために、動画操作機能や遅延を取り戻すための早送りを設けた。白石らの研究での実験条件では、1 グループ1名を 5 グループで行っており、投票は行っていなかった。この時、日本語対応手話の講演者の「講演者の内容」の場合、通訳品質は 5.9 である。本研究では、2回目日本語対応手話の講演者の時、通訳品質は 5.8 であった。中間手話の講演者の「講演者の内容」の場合、白石らの研究では、通訳品質は 6.6 である。本研究では、1, 3, 4, 5 回目の中間手話の講演者の時、通訳品質は最小 6.0、最大 6.6 であった。全体的に通訳品質の比較を行うと、白石らの研究と同等かそれ以下であった。しかし、白石らの研究で 1 グループ3名以上行った場合は、通訳品質は高くなることが想定できる。その代わりに、17名以上必要である。本研究では最大11名で同等かそれ以下の通訳品質で実現できることが分かった。

中間手話の時の平均通訳品質は 6 以上であったが、日本語対応手話の時は 6 未満であった。ログを分析すると、2回目の実験ではルールの不理解により、前の通訳区間のみ手話文字通訳を行うことが頻繁であったため、通訳品質が大きく下がってしまった。

通訳品質の各評価者の評価結果に対して、ケンドール一致係数を用いて一致度の評価を行った。結果、3回目から5回目の間は評価者の一致度がかなり低く、有意差が認められなかった。

つまり、評価者のばらつきが大きく、人数不足であった。原因は明確ではないが、手話歴を聞いたところ、1年8ヶ月の評価者がいたことが分かった。単純に手話読み取りに慣れていなかったことが考えられる。他にも考えられる原因として、評価者は手話講演者の手話と口が合わない時や、手話表現が曖昧の時は読み取りにくくなるため、通訳品質の評価が困難となった可能性もある。今後は、事前に、評価者の手話や日本語に関する能力を調査して揃えた上で、評価を実施する必要があるだろう。また、通訳品質の評価の際に、各通訳区間の手話の読み取りやすさを同時に評価してもらう方法も検討したい。

1つの通訳区間に対し、1名のワーカでも白石らの研究の通訳品質と同等かそれ以下の通訳品質を出せた理由は、動画操作機能を設けたことが理由だと考えられる。動画操作機能の主観評価結果のコメントを見ると、担当する通訳区間の再生位置の終わりになると、自動で巻き戻しされるため、再生速度を1.0倍速にする者もいた。つまり、再確認を前提に通常の再生速度で動画を見たものがいたことが分かる。他にも、デフォルトの再生速度がちょうど良いと答える者もあり、再生速度が速すぎても分からない、遅すぎても入力しにくいことが分かった。ワーカによっては再生速度を落とした方が手話を読み取りやすい者がいたことが分かった。しかし、今回は主観評価のみであり、動画操作機能が直接的に通訳品質に影響を与えたかは自明ではない。今後は、動画操作機能ありとなしでの比較実験を検討し、どの動画操作機能が通訳品質に影響を与えるか評価を行う。

通訳品質は5.8から6.6の間であったが、誤りはまだ残っている。1つの通訳区間に対し、1名のワーカで行う方法では、通訳品質の低い通訳結果に対して何もできない課題がある。受講者にとって、誤りのある通訳結果が字幕に表示されるだけで内容理解が困難になることが考えられる。更なる改善のために、修正タスクを設け、修正ワーカが文章の誤りを修正する方法を検討する。

5.10.7. 連係入力

4章の予備実験では訓練1回のみであり、連係入力を行った者は1名のみであった。本実験では手話文字通訳や連係入力に慣れてもらうために、訓練を4回行ってから実験用動画の手話文字通訳を行ってもらった。結果、4章では欠落率と重複率と共に34%であったものが、今回では欠落率の平均は19%(SD 7%)、重複率の平均は6%(SD 2%)であった。アンケートで前後の入力文を確認したか聞いた結果、確認しなかったワーカは2回目実験で1名のみであり、確認しない時があったと答えたワーカは各実験で1名から2名であり、それ以外の者は毎回確認していた。毎回確認したワーカらのコメントを見ると、できるだけ情報を正確に提供したい、文脈は大丈夫か確認した、余裕があったからといった回答があった。つまり、字幕を読む受講者のために、前後の繋がりや文脈に間違いがないか確認を行ったワーカがいたため、欠落率や重複率が下がったと考えられる。また、訓練によって連係入力を行う余裕が生まれたことも考えられる。確認しない時があったワーカらのコメントを見ると、連係入力は意識していても時々忘れることがある者や、入力が遅かったため確認しない時があった者がいた。ようするに、連係入力を時々忘れることや、入力に時間がかかりすぎて余裕がない時、重複や欠落を残したままタスク完了することがあると分かった。確認しなかったワーカらのコメントを見ると、ついていくのに精一杯だったからといった回答があった。つまり、手話

文字通訳に精一杯で連係入力を行う余裕がなかったと考えられる。確認しないワーカや確認しない時があったワーカからのコメントから、連係入力のみでは通訳区間の間の文の欠落や重複を完全に無くすことは厳しいことが分かった。今後は、修正タスクを設け、修正を担当するワーカが通訳を担当するワーカの入力結果を見て、通訳区間の間の文の重複や欠落を修正する方法を検討する。

5.10.8. 通訳区間の間の文の欠落と重複の発生理由

各実験の通訳区間の間の文の欠落率と通訳品質を見ると、1回目と2回目と3回目では欠落率が20%以上であり、通訳品質は1回目と2回目では約6.0以下であり、3回目のみ6.4と平均値に近い値であった。逆に、4回目と5回目では欠落率が約12%以下であり、通訳品質は6.6以上であった。1回目と2回目、4回目と5回目の欠落率と通訳品質を見ると、通訳品質が高いほど欠落率は低くなる傾向にあることが考えられる。おそらく、通訳区間の間に分からない手話があった時、その手話を日本語に通訳できないことが原因で、通訳品質が下がったことが考えられる。しかし、3回目実験では通訳品質が1回目と2回目より高いにも関わらず、欠落率が一番高かった。偶然的に通訳区間の間に分からない手話が良く出たなども考えられるが、今後は他に何らかの原因があるか分析を検討する。

通訳区間の間の重複率の平均は5.7(SD 1.9)であり、欠落率の平均より低かった。おそらく、通訳区間の間の文の重複は欠落と比べて、重複の文がほぼ同じであればすぐ気付きやすいからではないかと考えられる。しかし、依然として残っていた理由は、連係入力の忘れや、手話表現は同じだが訳し違いによって重複部分であることに気づけなかったことが考えられる。ログを分析すると、通訳区間の間の重複部分の手話文字通訳結果が「フォント」と「フォント」であり、同じ通訳結果であるにも関わらず重複が残っていたため、確認忘れが原因だろう。また、「考察」と「参考」もあり、同じ手話表現であるが通訳の結果が異なる場合もあったため、「参考」を入力したワーカは「考察」が同じ手話表現の通訳結果であることに気づけなかったことが考えられる。

各実験の通訳区間の間の重複率を見比べると、1回目と2回目では6.2以上であり、3回目と4回目と5回目では5.1%以下であり、3回目から低くなっていた。各実験の通訳品質も同様、3回目から全体の平均値より高くなっている。理由は、5.3節で述べた通り、1回目と2回目実験では、一つ前の通訳区間の手話文字通訳を行うワーカがいたため、より多くの文の重複が発生してしまったことが考えられる。今後は、入力開始時間や終了時間の制御を行い、確実にタスク開始時に入力を促すようなインタフェースを検討する。

5.10.9. 再生倍率 1.5 倍速で遅延を取り戻す方法の主観評価結果

手話文字通訳のタスクを終えた後、遅延が発生するため、再生速度を1.5倍速にしてリアルタイムに追いつく方法を検討した。予備実験では、全員がちょうどいいと答えていた。しかし、文献[18]では、手話習熟度や手話コンテンツにより読み取りやすい再生速度が異なるとされている。そこで、実験を5回行ない、複数の講演者の動画に対して再生速度1.5倍率についてどう感じたか回答してもらった。結果、全ての実験でほとんどのワーカがちょうどいいと答えていたが、速く感じたワーカ

やや遅く感じたワーカーもいた。やや速く感じたワーカーや速く感じたワーカーのコメントを見ると、少し速くてほとんどの手話が読み取れず所々で読み取ったことや、手話が読み取れなかったことが分かった。やや遅く感じた者のコメントでは、非常に遅れた時でも 1.5 倍速だから焦ったことや、講演者の手話の速さが遅くて合間も多かったからなどがあつた。つまり、早送りが速く感じる者は手話が読み取りにくいこと、遅く感じる者は長く待つことがストレスを感じるということが分かった。手話が遅く感じるワーカーにとって再生速度をあげても良い考えを持つワーカーがいることが分かった。他にも、入力完了後の遅延が大きく、次のタスクがなかなか割り当てられず、ストレスを感じるワーカーがいることも分かった。手話話速を分析すると、3回目から5回目の講演者は1回目や2回目の講演者より遅かった。2回目で速く感じたワーカーは3回目ではちょうど良いと回答しており、おそらく手話話速に影響あるからだと考えられる。今回は早送りの再生速度をデフォルトで 1.5 倍速としたが、手話話者の手話話速を基準にデフォルト値を決める方法も考えられる。ワーカーが手動で早送りの再生倍率を調整する方法も考えられるが、再生倍率が低いほど、遅延を取り戻す時間が長くなるため、必要人数が増えてしまう。ワーカーがタスクをより早く割り当てられるために再生倍率をかなり高くした場合、手話読み取りが困難で内容理解に影響を与えることが考えられる。今後は、動画解析で手話話速を求めて早送りの再生速度のデフォルト値を変更する方法や、早送りの再生速度を手動で手話の読み取れる速度に調整するように促すインタフェースを検討する。

5. 10. 10. 通訳回数と通訳の楽しさ

クラウドソーシングを用いて手話文字通訳を実現する場合、ボランティアのワーカーが継続して利用してもらうにはタスクを楽しめることが重要だと考えられる。しかし、予備実験では、ワーカー人数が4名のため、通訳回数が最大だったワーカーは楽しくないと回答していた。逆に通訳時間が平均的に長かったワーカーはタスク完了後休憩できるため、やや楽しかったと答えていた。そこで、必要人数を推定し、今回の実験では必要人数に近いワーカー数で実験を行い、通訳の楽しさの主観評価を行った。結果、全ての実験でほとんどのワーカーが楽しさを感じていた。楽しくなかったワーカーは2回目実験で1名だけであり、他の実験では0名であった。予備実験よりワーカー数を増やしたため、タスク完了後に待機できる時間が長くなったことが一つの要因だと考えられる。

楽しかったと回答したワーカーからのコメントを見ると、「助けられる側ではなく、助ける側に回れたことが心嬉しく感じる」と回答したワーカーや、通訳が好きなワーカーがいることが分かった。しかし、どちらでもないと答えたワーカーもいた。コメントを見ると、手話を読み取りにくかったことが原因であった。つまり、講演者の手話を読み取りにくいことがストレスとなったことが考えられる。他にも、やや楽しくなかったと回答したワーカーが1名いた。コメントを見ると、とても疲れたことが分かった。ログを分析すると、そのワーカーは全ての実験の中で通訳回数が最大であり、27回と2回目の平均の2倍以上であった。5.3節で述べた通り、一つ前の通訳区間の手話文字通訳を行っていたワーカーのため、6秒未満でタスクを完了させることが頻繁であった。また、再生速度が速く感じた者でもあり、早送り中の手話を読み取れなかったと回答していた。これらの複合的な理由により、他のワーカーと比べて負担がかなり大きかったことが考えられる。3回目以降ではルールを理解し、通訳回数が減ったため、ど

ちらでもいいと回答するようになっていた。必要人数に近いワーカ数で行えば、予備実験のように特定のワーカに負担が大きくなるのではなく、負担を分散することができ、通訳に楽しさを感じやすくなることが分かった。

5. 10. 11. SUS を用いた主観評価結果

ボランティアが継続利用してもらうためには、通訳の楽しさのみならず、システムの使いやすさや理解しやすさも重要となる。しかし、4章の予備実験では、通訳画面のユーザビリティの評価を行っていなかった。そこで、本実験では、手話文字通訳を終えた後、SUSを用いた主観評価を行なった。結果、ユーザビリティがやや悪いことが分かった。各質問項目の回答結果を見ると、以下の質問項目に対して肯定的な傾向であった。

- ・「(1)このウェブサイトをしばしば利用したいと思う。」
- ・「(5)このウェブサイトにあるコンテンツ(動画、指示文、入力枠、動画操作機能など)は十分に統一感があると感じた。」
- ・「(6)このウェブサイトでは一貫性のないところが多々あったと感じた。」
- ・「(9)このウェブサイトを利用できる自信がある」

つまり、ワーカは通訳画面の機能には統一感があると感じ、利用できる自信があり、継続の意欲を示していた。この結果から、ボランティアでも自信を持ちつつ、継続利用してもらえる可能性があることが分かった。しかし、「(10)このウェブサイトを利用し始める前に知っておくべきことが多くあると思う。」は否定的な傾向であった。つまり、ワーカへの指示文、遅延を取り戻すための早送り、連係入力など、知っておくことが多かったことが考えられる。2回目の実験でも誤解を与えたワーカがいたのも、知っておくことが多かったことが1つの要因だと考えられる。今後は、通訳画面の体験と説明を併用したチュートリアル画面を設け、SUSを検討する。

5. 10. 12. まとめ

本実験では、提案するシステムの有効性を確認するために、複数人の講演者に対し、必要人数に近いワーカ数で実験を行なった。白石らの研究では、カウンタ係と通訳ワーカ、投票ワーカを含めて17名以上要するため、人的コストが高い。また、カウンタ係の負担や、通訳区間の長さがまちまちである。本研究では、通訳区間を自動で、一定時間で区切り、重複し、文の欠落や重複に対して連係入力で解決できるか検討した。通訳品質の保障のために講演者はライブ動画配信を行い、通訳画面に動画操作機能を設け、映像の見直しや再生速度を遅くすることを可能とした。さらに、動画操作機能を用いたときに発生する遅延を取り戻すために早送りを設けた。この方法で、1つの通訳区間に対し、1名のワーカで行い、通訳品質を保ちつつ、必要人数を減らせるか検討した。結果、必要人数は8名から11名のワーカを要することが分かった。本研究の方が少ない人数で手話文字通訳を実現できることが分かった。しかし、必要人数を満たしても通訳時間のばらつきが大きいため、遅延が発生し、字幕の遅れが徐々に大きくなっていた。通訳品質において、白石らの研究より通訳品質は低かったものの、平均的に80%から100%の間の通訳結果を表出することができ

た。しかし、通訳品質の主観評価を行なった評価者のばらつきが大きく、通訳品質の評価者の募集条件の見直しが必要という課題も見つかった。

予備実験では、通訳区間9秒であり、平均通訳時間は白石らの研究の通訳時間より長かった。字幕の遅れが大きいと受講者は講演や周囲の状況把握を困難となる。本実験では、通訳区間を6秒にした結果、白石らの研究の通訳時間より字幕の遅れを減らすことができた。

予備実験では、連係入力を行なったワーカが1名のみであり、訓練不足だと考えた。本実験では、手話文字通訳と連係入力の慣れの主観評価を行った結果、約20分の訓練を要し、実験用動画では、主観評価ではほとんどのワーカが連係入力を行っていた。いちご一会とちぎ国体[24]ではボランティアが情報支援スタッフの研修を1回行うのに2時間要するのに対し、本システムでは訓練時間約20分で主観的に慣れることができる。

通訳区間の間の文の欠落や重複に対して、連係入力で解決できるか検討した結果、ほとんどのワーカが連係入力を主観的に行っており、欠落率は6%、重複率は19%であった。予備実験ではどれも34%であったため、本実験では予備実験と比べて欠落や重複が減少できたが、依然として残っている。

予備実験では、通訳回数が最大で楽しくなかったワーカがいたため、推定された必要人数に近いワーカ数で本実験を行なった。結果、ほとんどのワーカが通訳に楽しさを感じていた。SUSの主観評価を行なった結果、システムを利用できる自信や、継続利用に関して肯定的であった。本システムはボランティアでもタスクを楽しむことができ、継続利用できる可能性がある。しかし、システムについて知っておくべきことが多くあると感じたワーカが多くいることが分かった。1名のワーカはルールの誤解により通訳回数が最大となり、楽しくなかった。今後は、通訳画面の体験と説明を併用したチュートリアル画面を設けることで、システムを理解しやすくなるか検討する。また、入力開始時間や終了時間の制御を行い、確実にタスク開始時に入力を促すようなインタフェースを検討する。

複数の講演者に対して、複数のワーカが早送り再生速度1.5倍速は速く感じるか、主観評価を行なった。結果、ほとんどのワーカがちょうど良いと感じていた。しかし、速く感じたワーカは手話の読み取りが困難であった。

これらの結果から、システムを改善するために、必要人数を事前に決められるようにするために通訳タスクに制限時間を設ける。さらに、誤りや欠落と重複をなくすための修正タスクを設ける。また、早送り再生速度が合わなく感じるワーカのために、手動で調整する方法や、手話話速に応じてデフォルト再生速度を決める方法を検討する。そして、通訳品質の評価者の募集条件の見直しを行い、通訳品質の主観評価と一致度の評価を行う。

第6章 結論

6.1. まとめ

本研究では、手話が読める人たちがボランティアとして手話を文字に変換する方法を実現するために、クラウドソーシングを用いた手話文字通訳システムの提案とその有効性の評価を行った。本システムは、講演者の手話をライブ動画として配信し、手話動画を短く区切ることで通訳区間を設定し、それらの通訳区間を複数人のワーカに割り当てて手話を文字化(手話文字通訳)するタスクを担当してもらう。タスクが割り当てられた1名のワーカは通訳区間の手話動画を見ながら手話文字通訳を行う。最後に、各ワーカの入力結果を通訳区間の順に連結して字幕として表示する。全てのワーカはキューで管理され、システムはワーカがタスクを終了した際に改めてキューに挿入し直すことで、繰り返しタスクを担当させていくことが可能になり、手話文字通訳を実現できる。字幕表出までの遅延やワーカの人的コストを減らすこと、結果として出力される字幕の質を維持するために、通訳区間を一定の時間として自動で区切る方法を採用し、1名のワーカで1通訳区間を担当できるように、映像の見直しや再生速度調整などの動画制御機能を設けた。また、通訳区間の間の文の欠落や重複を減らすために、ワーカ間での連係入力について検討した。さらに、タスク完了後の動画再生位置の遅延を取り戻すために、現在の再生位置まで早送りで追いつかせる方法を導入し検討した。

本研究で実装したシステムを使用して、予備実験と本実験の2回の実験を通して有効性の評価を行なった。本システムが字幕表示の遅れや誤り、人的コストを減らせたか、先行研究(白石らのシステム[10])と比較を行なった。その結果、本システムでの手話文字通訳の実現に必要なワーカ数は8名から11名の人数であり、先行研究(17名以上とされている)よりも必要ワーカ数を減らすことができた。字幕表示の遅れについては、タスク開始からタスク終了までの時間が17.0秒程度であり、先行研究よりも減らすことができた。通訳品質の主観評価結果は先行研究と同等かそれ以下であった。本実験の5回の手話文字通訳の実施において、部分的に通訳品質の低い通訳区間が見られたが、平均的に80%以上100%未満の通訳結果を表出できていた。

受講者(字幕利用者)のために、通訳区間の間の文の欠落や重複をなくす連係入力を検討し、文の欠落や重複を減らせたか客観評価を行なった。その結果、ワーカ数が不足していた予備実験においては、欠落率と重複率は共に33%であった。システムに必要なワーカ数を推定し、ワーカに訓練をもらった上で実施した本実験では、欠落率が6%、重複率が19%となり改善がみられた。しかし、受講者の内容理解や字幕の読みやすさを考慮すると、現状の欠落・重複率では十分ではない可能性があるため、さらなる改善方法が求められる。

入力完了後、遅延を取り戻すための再生速度1.5倍速は速く感じるか主観評価を行なった。その結果、遅く感じていた、あるいは速く感じていたワーカが1割程度いたが、予備実験でも本実験でもほとんどのワーカがちょうど良いと回答していた。1.5倍速が速く感じていたワーカは、手話の読み取りが困難という理由であった。

経験の少ないボランティアの参加意欲を高めるために、短時間の訓練で参加できつつ、手話文字通訳を楽しめることを評価した。その結果、初めてのワーカが手話文字通訳や連係入力に慣れるのに要する時間は20分程度であり、実用上問題とならない程度の短い時間で主観的に慣れることが分かった。さらに、約9割のワーカが通訳を楽しむことができた。SUSの結果から、システム利用前に知るべきことが多いことが示されたが、システムのインタフェースの一貫性や継続的な利用については肯定的であることが明らかになった。

本システムは、先行研究のシステムより字幕の遅れを減らしつつ、少ない人数で手話文字通訳を実現できた。通訳区間の間の文の重複や欠落や、通訳品質についての課題が明らかになったが、ほとんどのワーカが短い訓練時間で参加を可能にしつつ、通訳を楽しむことができ、継続的な利用に肯定的であった。これまでは、手話で講演が行われる場合に、手話の分からない者のために手話通訳と文字通訳を依頼する必要があった。本研究で提案し有効性を評価したシステムは、今後も研究開発を重ねることで、ろう者・難聴者の手話が読める専門家以外の者だけで手話文字通訳を行うことができる手段の1つになることが期待できる。また、本実験に参加したワーカから得られた「助けられる側ではなく、助ける側に回れたことが心嬉しく感じる」というコメントからも、従来は支援される側であったろう者・難聴者が支援する側に回り、手話講演者や手話の分からない者を支援することができる新たな経験を提供することが可能になる。本研究が、聴覚障害者の情報保障に関わる環境を、より改善していく一助となることを期待したい。

6.2. 今後の課題

手話文字通訳システムを実現する上での課題と今後の課題を述べる。1つの通訳区間に対し、1名のワーカでは、誤りや通訳区間の間の文の重複と欠落が依然として残ることが分かった。また、必要人数を満たしていたとしても、通訳時間のばらつきが大きいと、遅延が大きくなることがあった。今後は、必要人数の増減を防ぐために通訳タスクに制限時間を設けるなどの工夫が考えられる。さらに、通訳を担当するワーカの入力結果の誤りを修正する修正タスクを設け、修正を担当するワーカが誤りや通訳区間の間の文の欠落と重複を修正する方法も考えられる。

予備実験では、タイピング速度が速いほど通訳時間が短くなる傾向が見られた。つまり、タイピング速度が遅いワーカは通訳区間の時間が長いと、入力に時間がかかり、字幕の遅れが大きくなる。今後は、各ワーカに適した通訳区間の時間を動的に調整する方法の検討が求められる。

通訳品質の評価の際、評価者のばらつきが大きく見られた。原因は手話講演者の手話の読み取りやすさや、評価者の手話や日本語に関する能力の違いなどが考えられる。今後は、事前に、評価者の手話や日本語に関する能力を調査して揃えた上で、評価を実施する必要があるだろう。また、通訳品質の評価の際に、各通訳区間の手話の読み取りやすさを同時に評価してもらう方法も検討したい。

遅延を取り戻すための再生速度(再生倍率)について約1割のワーカは速度が速い／遅いと感じていた。また、講演者の手話技能や手話話速も強く影響することが考えられる。今後は、ワーカ

が早送りの再生倍率を手動で調整する方法, または動画解析を用いて手話話速に適した再生倍率に調整する方法を検討したい.

SUS の結果では, ワーカは手話文字通訳のタスクを楽しめる傾向, 継続する意欲を示す傾向にあったが, 初めてのワーカにとって手話文字通訳システムについて知っておくべきことが多いことが明らかになった. また, 前の通訳区間のみ手話文字通訳を行ったワーカがおり, 通訳を楽しめなかったとの回答もあった. 今後は, 説明と体験を併用できるチュートリアル画面を作成すると共に, 入力開始時間や終了時間の制御を行い, 確実にタスク開始時に入力を促すようなインタフェースを検討する必要があるだろう.

謝辞

本研究に関して終始ご指導ご鞭撻をいただきました若月大輔教授と皆川洋喜教授に深く感謝致します。また、貴重な時間を割いて本論文をご精読頂き有用なコメントをいただきました修士論文主査の加藤伸子教授と副主査の平賀瑠美教授に深く感謝致します。研究について活発な議論をさせて頂いた若月教授、皆川教授と共に研究活動を行なっている 4 年生に深く感謝致します。最後になりましたが、修士課程に進学する機会を与えてくださり、ありとあらゆる場面で私を暖かく見守り続けてくれた両親に深く感謝致します。

参考文献

- [1] 田中 康平, 若月 大輔, 皆川 洋喜, “クラウドソーシングによる手話文字通訳のためのタスクの基礎的検討”, 信学技報, vol. 118, no. 491, WIT2018-83, pp. 165-170, 2019 年 3 月.
- [2] Preminger, J.E., Levitt, J.E.: Computer-Assisted Remote Transcription (CART): A Tool to Aid People Who Are Deaf or Hard of Hearing in the Workplace. *Volta Review*, Vol.99, No.4, pp.219-230(1997)
- [3] Elliot, L.B., Stinson, M.S., McKee, B.G., Everhart, V.S., Francis, P.J.: College Students' Perceptions of the C-Print Speech-to-Text Transcription System. *Journal of Deaf Studies and Deaf Education*, Vol.6, No.4, pp.285-298(2001)
- [4] Kurita, S.: Development of CART software “IPtalk” and captioning services via personal computers: Communication support for the hearing-impaired. *Journal of Information Processing and Management*, Vol.59, No.6, pp.366-376(2016) <https://doi.org/10.1241/johokanri.59.366>
- [5] Wakatsuki, D., Kato, N., Shionome, T., Kawano, S., Nishioka, T., Naito, I.: Development of Web-Based Remote Speech-to-Text Interpretation System captiOnline. *JACIII*, Vol.21, No.2, pp.310-320(2017) <https://doi.org/10.20965/jaciii.2017.p0310>
- [6] 白澤麻弓, 三好茂樹, 石野麻衣子, 吉川あゆみ, 松崎丈, 中野聡子, 岡田孝和, 太田晴康, “パソコンノートテイク における連係入力のプロセス分析,” 日本特殊教育学会 第 48 回大会発表論文集, pp239, 2010.
- [7] 栗田茂明. “パソコン要約筆記における連係入力方法の分析と Q 方式の提案”. 画像電子学会 年次大会予稿集. 画像電子 学会. 2012. <http://www.nck.or.jp/shiryoku/120707Q-method.pdf>, (accessed 2021-08-14).
- [8] Lasecki, W.S., Miller, C.D., Sadilek, A., Abumoussa, A., Borrello, D., Kushalnagar, R., Bigham, J.P.: [Real-time Captioning by Groups of Non-Experts](#). In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST 2012)*, pp.23-34(2012)
- [9] Lasecki, W.S., Miller, C.D., Bigham, J.P.: [Warping Time for More Effective Real-Time Crowdsourcing](#). In *Proceedings of the International ACM Conference on Human Factors in Computing Systems (CHI 2013)*, pp.2033-2036(2013)

- [10] Shiraishi, Y., Zhang, J., Wakatsuki, D., Kumai, K., Morishima, A.: Crowdsourced Real-Time Captioning of Sign Language by Deaf and Hard-of-Hearing People. International Journal of Pervasive Computing and Communications, Vol. 13 No. 1, pp. 2-25(2017)
- [11] Kumai, K., Zhang, J., Shiraishi, Y., Wakatsuki, D., Kitagawa, H., Morishima, A.: Group Rotation Management in Real-Time Crowdsourcing. Proceedings of 19th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2017), pp.23-31(2017)
- [12] Kumai, K., Matsubara, M., Shiraishi, Y., Wakatsuki, D., Zhang, J., Shionome, T., Kitagawa, H., Morishima, A.: Skill-and-Stress-Aware Assignment of Crowd-Worker Groups to Task Streams. The sixth AAAI Conference on Human Computation and Crowdsourcing (HCOMP2018), pp.88-97(2018)
- [13] Shionome, T., Hashimoto, H., Zhang, J., Shiraishi, Y., Wakatsuki, D., Seki, Y., Morishima, A.: Complement of Incomplete Task Results for Real-time Crowdsourcing Interpretation. Proceedings of 21th International Conference on Asian Language Processing (IALP 2017), pp.359-362(2017)
- [14] Hashimoto, H., Matsubara, M., Shiraishi, Y., Wakatsuki, D., Zhang J., Morishima, A.: A Task Assignment Method Considering Inclusiveness and Activity Degree. Proceedings of The Second IEEE Workshop on Human-in-the-loop Methods and Human Machine Collaboration in BigData (IEEE HMDData2018), pp. 3498-3503(2018)
- [15] YouTube Live: <https://www.youtube.com/live?gl=JP&hl=ja>, 2020/1/2
- [16] Node.js: <https://nodejs.org/>, 2020/1/2
- [17] YouTube Player API: https://developers.google.com/youtube/iframe_api_reference?hl=ja, 2020/1/2
- [18] 日本語入力スピード測定, <https://flickromaji.oka-ryunoske.work/>, 2020/1/2
- [19] Google Chrome, https://www.google.com/intl/ja_jp/chrome/, 2020/1/2

- [20] 若月 大輔, 加藤 伸子, 村上 裕史, 皆川 洋喜, 西岡 知之, 河野 純大, 内藤 一郎, 三好 茂樹, 石原 保志, ”断続的な停止をともなう手話映像の内容理解に関する評価”, 電子情報通信学会 HCG シンポジウム 2013
- [21] 滝口 雄介, 大澤 洋, 阿部 剛, 平林 大輔, 奥 寛昭, 磯野 春雄, ”聴覚障害者に対する手話ニュース映像の好ましい再生速度”, 2005 年映像情報メディア学会冬季大会(ITE Winter Annual Convention 2005)
- [22] John. System usability scale (SUS): a quick-and-dirty method of system evaluation user information. DEC, 1986.
- [23] How to Measure Product Usability with the System Usability Scale (SUS) Score, <https://uxplanet.org/how-to-measure-product-usability-with-the-system-usability-scale-sus-score-69f3875b858f>, 2020/1/8
- [24] 研修(情報支援スタッフ養成講座)について,
<https://www.tochigikokutai2022.jp/volunteer/informationvolunteer/>研修(情報支援スタッフ養成講座)について/, 2021/8/17

本研究に関する成果・発表等

- [A1] 田中康平, 若月大輔, 皆川洋喜, クラウドソーシングに基づく手話文字通訳のタスクのための研究, 第 14 回日本聴覚障害学生高等教育支援シンポジウム, 2018 年 10 月.
- [A2] 田中 康平, 若月 大輔, 皆川 洋喜, ”クラウドソーシングによる手話文字通訳のためのタスクの基礎的検討”, 信学技報, vol. 118, no. 491, WIT2018-83, pp. 165-170, 2019 年 3 月.
- [A3] 田中康平, 若月大輔, 皆川洋喜, ”クラウドソーシングによる手話文字通訳”, 第 15 回日本聴覚障害学生高等教育支援シンポジウム, 2019 年 11 月.
- [A4] 田中 康平, 若月 大輔, 皆川 洋喜, “クラウドソーシングを用いた手話文字通訳 ～ ライブ動画配信を活用したシステムの試作 ～”, 信学技報, vol. 119, no. 322, WIT2019-39, pp.95-100, 2019 年 12 月.
- [A5] Kohei Tanaka, Daisuke Wakatsuki, Hiroki Minagawa, “A Study Examining a Real-Time Sign Language-to-Text Interpretation System Using Crowdsourcing”, ICCHP2020: Computers Helping People with Special Needs pp 186-194, 2020 年 9 月.
- [A6] 若月 大輔, 田中 康平, 皆川 洋喜, “クラウドソーシングによる手話文字通訳に関する研究”, 筑波技術大学テクノレポート, Vol.28, No.1, pp.61-62, 2020 年 12 月.

付録

A1 開発環境

名称/種類	用途	備考
Visual Studio Code	プログラム作成 デバッグ	
Node.js v15.4.0	サーバー配信 サーバープログラミング	サーバーサイドのJavascript
socket.io	クライアントとサーバーとのリアルタイム処理を可能にするモデル	Node.jsのnpmパッケージによりインストール可能
さくらのレンタルサーバー	Node.jsを用いたサーバー公開	

A2 第5章の本実験のアンケート

A2.1 手話文字通訳システムの慣れについてのアンケート

練習1回目のアンケート

手話文字通訳システムを使用する際の慣れについて質問します。

学籍番号を入力してください。 （記述例：101010） *

記述式テキスト（短文回答）

手話文字通訳（手話を読み取りながら手話を日本語に通訳しながら文字を入力する作業）の慣れ *
はどうでしょうか？

慣れない

1

2

3

4

5

慣れた

理由を教えてください。 *

記述式テキスト（長文回答）

動画操作の時、以下のどれかを使用しましたか？（両方の場合は両方チェックを入れてください） *

☐ ボタン

☐ ショートカットキー

ボタンまたはショートカットキーによる動画操作（1秒戻る、再生速度調整、再生停止）の慣れはどうか？ *

	1	2	3	4	5	
慣れない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	慣れた

理由を教えてください。 *

記述式テキスト（長文回答）

前後の入力文に入力漏れや重複があるか確認することの慣れはどうか？ *

	1	2	3	4	5	
慣れない	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	慣れた

理由を教えてください。 *

記述式テキスト（長文回答）

A2.2 手話文字通訳実験についてのアンケート

手話文字通訳実験のアンケート PC版

実験協力ありがとうございました。
このアンケートは研究システムの改善，収集したデータをより詳細に分析するために実施しています。
アンケートは10分程度で回答できます。頂いた回答はアンケートの目的以外には一切使用いたしませんので、率直なご感想をお聞かせください。
どうぞ宜しくお願いします。

学籍番号を入力してください。 *

記述式テキスト（短文回答）

手話歴を大体で構いませんので教えてください（例：1年） *

記述式テキスト（短文回答）

手話文字通訳について

ここから手話文字通訳についていくつか質問します。

講演者の手話は分かりやすかったか *

1 2 3 4 5
分かりにくかった ○ ○ ○ ○ ○ 分かりやすかった

理由を教えてください *

記述式テキスト（長文回答）

講演内容は理解できましたか *

1 2 3 4 5
理解できなかった ○ ○ ○ ○ ○ 理解できた

理由を教えてください *

記述式テキスト（長文回答）

PCのOSはどれですか *

windows

Mac

ChromeOS

その他...

タッチタイプ（キーボードを見ないで入力）はできますか *

はい

いいえ

この中で使った機能はどれですか（複数可） *

再生倍率調整

1秒戻る

再生・停止

特になし

再生倍率調整を使った方、理由を教えてください。使わなかった方はデフォルトである0.8倍速についてどう感じたか教えてください。 *

記述式テキスト（長文回答）

1秒戻るを使った方、理由を教えてください。使わなかった方は「特になし」と入力してください。 *

記述式テキスト（長文回答）

再生・停止を使った方、理由を教えてください。使わなかった方は「特になし」と入力してください。 *

記述式テキスト（長文回答）

通訳後、遅れを取り戻すために再生倍率を1.5倍速にしましたが、早く感じましたか。（3は「ちょうど良い」です） *

	1	2	3	4	5	
早い	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	遅い

理由を教えてください。 *

記述式テキスト（長文回答）

通訳中に前後の入力文から入力漏れや重複があるか確認しましたか？ *

- ☐ 毎回確認した
- ☐ 確認しない時があった
- ☐ 確認しなかった

理由を教えてください。 *

記述式テキスト（長文回答）

通訳は楽しかったですか？ *

	1	2	3	4	5	
楽しくなかった	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	楽しかった

理由を教えてください *

記述式テキスト（長文回答）

通訳は楽しかったですか *

	1	2	3	4	5	
楽しなかった	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	楽しかった

理由を教えてください *

記述式テキスト（長文回答）

自由記述（不便な面や改善して欲しい機能などの意見、感想など、なんでも構いません。）

記述式テキスト（長文回答）