

筑波大学大学院博士課程
理工情報生命学院
システム情報工学研究群修士論文

多人数ビデオ会議における話者交替のための
参与役割提示手法

飯塚 陸斗
修士（工学）
（情報理工学位プログラム）

指導教員 志築 文太郎

2023年3月

概要

多人数ビデオ会議において、円滑な話者交替を行うことは、対面会議と比較して困難である。これは多人数ビデオ会議において、話し手や受け手といった会話における各参加者の役割（参与役割）を把握することが困難であることに起因すると考えられる。そこで本研究では、多人数ビデオ会議において参与役割の把握を支援し、話者交替を促進するシステムを提案する。これまでに発話欲求が表出する発話予備動作から次に話しそうな参加者をハイライト表示する研究や頭部動作から受け手を予測する研究がされている。また、著者による先行研究として、ビデオ会議における視線配布から参与役割を推定し、ビデオウィンドウのサイズを用いて提示する手法を提案している。しかし、それらの研究は参与役割の推定する精度が不十分であることや、発話欲求の高い参加者が複数人いた際に話者交替を促進できないと考えられる。そこで本研究では複数の非言語情報から4つの参与役割を推定する手法を提案する。本研究では、参与役割のうち、発話の有無から「話し手」を推定する。次に発話欲求の高い参加者のうち、話し手の視線配布が向いている参加者を「受け手」、向いていない参加者を「傍参加者」として推定する。また、推定した参与役割を、ウィンドウサイズおよび配置を用いて提示する手法を提案する。本研究では、この手法を用いて話者交替が促進されたかを評価するために実験を行った。実験の結果、提案手法が参与役割の把握を促し、話者交替が促進する可能性があることが明らかになった。

目次

第1章	はじめに	1
1.1	本研究で扱う多人数ビデオ会議の定義	1
1.2	多人数ビデオ会議における課題	2
1.3	課題に対する先行研究	2
1.4	目的とアプローチ	3
1.5	貢献	3
1.6	本論文の構成	3
第2章	関連研究	4
2.1	多人数ビデオ会議における社会学的知見	4
2.1.1	会話に使用される非言語情報	4
2.1.2	話者交替について	5
2.1.3	参与役割とF陣形について	5
2.2	ビデオ会議を支援する研究	8
2.2.1	現在普及しているビデオ会議システム	8
2.2.2	非言語情報を共有することで支援する研究	8
2.2.3	非言語情報を参与役割として表示する研究	9
2.2.4	著者による先行研究	10
2.3	本研究の位置付け	10
第3章	設計指針	12
3.1	参与役割の推定手法	12
3.2	参与役割の提示手法	14
第4章	実装	15
4.1	ビデオ会議システムのシステム構成	15
4.2	非言語情報の検出	16
4.2.1	挙手の検出	16
4.2.2	開口の検出	17
4.2.3	頷きの検出	18
4.2.4	視線配布の検出	18
4.2.5	発話の検出	20

4.3	発話欲求度の算出	20
4.4	参与役割の推定	21
4.5	参与役割の提示	22
4.5.1	ウィンドウサイズを用いた提示	23
4.5.2	配置を用いた提示	23
第5章	実験	24
5.1	実験設計	24
5.1.1	実験条件	24
5.1.2	実験タスク	26
5.1.3	実験参加者	26
5.2	実験手順	27
5.3	仮説と評価項目	29
5.3.1	仮説1：参与役割提示条件では、話し手強調提示条件と比較して、参与 役割の把握が促進される.	29
5.3.2	仮説2：参与役割提示条件では話し手強調提示条件と比較して、円滑な 話者交替が促進される.	30
5.3.3	仮説3：参与役割提示条件では話し手強調提示条件と比較して、会議の 質が向上する.	30
第6章	実験結果および考察	31
6.1	会話分析	31
6.1.1	発話数	31
6.1.2	話者交替数	32
6.1.3	発話衝突の回数および確率	32
6.1.4	沈黙時間	34
6.1.5	参与役割の遷移	35
	- 話し手の前に受け手であったか	35
	- 受け手の後に話し手であったか	35
	- 傍参加者の後に傍観者であったか	36
	- 傍参加者の後に話し手であったか	37
	- 傍観者の後に話し手であったか	37
6.1.6	発話欲求度の平均	38
6.1.7	発話予備動作	39
	- 話し手への視線配布した時間	39
	- 頷き回数	40
	- 挙手した回数および時間	41
	- 開口した回数および時間	41
6.2	アンケートによる印象評価	42

6.2.1	参与役割の把握	42
	- 今自分が会話に参加していること/していないことを意識することができた (図 6.16)	42
	- 自分の話を聞いている参加者がわかった (図 6.17)	43
	- 自分がちゃんと聞いていることが相手に伝わっているように感じた (図 6.18)	44
	- 自分が話し終えるとき、次に話たさそうな人がわかった (図 6.19)	44
	- 自分が話し終えるとき、話したくない人がわかった (図 6.20)	45
	- 他の人が話し終えるとき、自分が話すべきかがわかった (図 6.21)	46
6.2.2	話者交替および会議への評価	46
	- 議論に集中できた (図 6.22)	46
	- 議論にストレスを感じた (図 6.23)	47
	- 議論に疲労を感じた (図 6.24)	48
	- 話者交替にストレスを感じた (図 6.25)	48
6.2.3	AUT による創造性の評価	49
	- アイデアの数はたくさん考えられたか (図 6.26)	49
	- 独創的なアイデアを考えることができたか (図 6.27)	49
	- 幅広い用途のアイデアを考えることができたか (図 6.28)	50
	- 詳細なアイデアを考えることができたか (図 6.29)	51
6.3	考察	51
6.3.1	仮説 1: 参与役割提示条件では、話し手強調提示条件と比較して、参与役割の把握が促進される.	51
6.3.2	仮説 2: 参与役割提示条件では話し手強調提示条件と比較して、円滑な話者交替が促進される.	53
6.3.3	仮説 3: 参与役割提示条件は話し手強調提示条件と比較して、会議の質が向上する.	54
第 7 章	今後の課題	55
7.1	参与役割の推定手法について	55
7.2	参与役割の提示手法について	55
7.3	実験設計について	56
第 8 章	おわりに	57
	謝辞	58
	参考文献	59
付録 A	実験におけるアンケート	66
A.1	条件間に行ったアンケート	67

A.2 全ての条件後に行ったアンケート	73
-------------------------------	----

目次

2.1	会話における参与役割.	6
2.2	話者交替時の視線のやり取り. 話し手と受け手は相互注視を行う. また, 傍参加者は話し手へ一方的に視線配布を行う.	7
2.3	F陣形における会話への参与の流れ. 会話へ途中から参与する人はまずR空間にとどまる(左図). その後P空間に身体を運び入れる(右図).	7
2.4	先行研究における提案手法. ウィンドウサイズを変化させる手法(右図). 視線配布を直接矢印としてビデオウィンドウに重畳表示する手法(左図).	10
3.1	参与役割の遷移. 実線の矢印は設計意図通りの遷移を示し, 破線の矢印は話者交替が失敗する可能性の高い遷移を表す.	13
4.1	本研究のシステム構成. 多人数ビデオ会議システムをWebアプリケーションとして実装した.	16
4.2	MediaPipeを用いた挙手の検出.	17
4.3	MediaPipeを用いた開口の検出.	17
4.4	MediaPipeを用いた頷きの検出.	18
4.5	MediaPipeを用いた顔向きの取得と視線配布の検出.	19
4.6	実装したシステム. 画面表示を用いて参与役割を提示する.	22
5.1	実験条件. 上: 参与役割提示(サイズ+配置)条件. 下左: 参与役割提示(サイズ). 下右: 話し手強調提示条件.	25
5.2	実験手順.	27
6.1	各条件における参加者の発話数の平均.	31
6.2	各条件における参加者の話者交替数の平均.	32
6.3	各条件における参加者の発話衝突の回数の平均. A: 自分が発話を開始した際に他の人が既に発話していた数. B: 自分が発話している際に他の人が発話を開始した数.	33
6.4	各条件における参加者の発話衝突の確率.	34
6.5	各条件における各グループの沈黙時間.	34
6.6	話し手の前に受け手であったかの評価.	35
6.7	受け手の後に話し手であったかの評価.	36
6.8	傍参加者の後に傍観者であったかの評価.	36

6.9	傍参加者の後に話し手であったかの評価	37
6.10	傍観者の後に話し手であったかの評価	38
6.11	各条件における参加者の発話欲求度の平均.	39
6.12	各条件における参加者の話し手へ視線配布した時間.	40
6.13	各条件における参加者の頷きを行った回数.	40
6.14	各条件における参加者の挙手した回数 (左) および時間 (右).	41
6.15	各条件における参加者の開口した回数 (左) および時間 (右).	42
6.16	各条件における b-1 「今自分が会話に参加していること/していないことを意識することができた」に対する回答.	43
6.17	各条件における b-2 「自分の話を聞いている参加者がわかった」に対する回答.	43
6.18	各条件における b-3 「自分がちゃんと聞いていることが相手に伝わっているように感じた」に対する回答.	44
6.19	各条件における b-4 「自分が話し終わるとき、次に話たさそうな人がわかった」に対する回答.	45
6.20	各条件における b-5 「自分が話し終わるとき、話したくない人がわかった」に対する回答.	45
6.21	各条件における b-6 「他の人が話し終わるとき、自分が話すべきかがわかった」に対する回答.	46
6.22	各条件における c-1 「議論に集中できた」に対する回答.	47
6.23	各条件における c-2 「議論にストレスを感じた」に対する回答.	47
6.24	各条件における c-3 「議論に疲労を感じた」に対する回答.	48
6.25	各条件における c-4 「話者交替にストレスを感じた」に対する回答.	48
6.26	各条件における a-1 「アイデアの数はたくさん考えられたか」に対する回答.	49
6.27	各条件における a-2 「独創的なアイデアを考えることができたか」に対する回答.	50
6.28	各条件における a-3 「幅広い用途のアイデアを考えることができたか」に対する回答.	50
6.29	各条件における a-4 「詳細なアイデアを考えることができたか」に対する回答.	51
6.30	各題材に対する綿密性の評価.	54

表目次

2.1	坊農ら [坊農 04] が述べた会話へ参与する手続き.	7
3.1	提案手法における非言語情報と参与役割の関係.	13
4.1	発話予備動作と発話欲求度.	21
4.2	ウィンドウサイズの大きさ.	23
4.3	ウィンドウを円形に配置する際の半径と参与役割.	23
5.1	各グループにおける条件および題材.	27
5.2	各条件実施後に参加者に回答させたアンケートを示す. AUT における創造性を評価 (a) するための4つの項目を10段階リッカート尺度を用いた回答欄にて回答させた. また直前まで行っていた会議における参与役割の把握 (b) および会議 (c) についての印象評価を5段階リッカート尺度を用いた回答欄にて回答させた.	28
5.3	全ての条件が終了した後に参加者に回答させたアンケートを示す. ここでは提案手法に対する印象を5段階で回答させた. またシステム, および実験に関する印象を記述式に回答させた.	29

第1章 はじめに

近年、ビデオ会議システムを用いた遠隔地間のコミュニケーションが普及している。ビデオ会議システムが普及した理由について、導入が容易であることおよび対面におけるコミュニケーションが減少したことが挙げられる [玉木 12]。ビデオ会議システムは、インターネットに接続できる環境およびスマートフォンまたはコンピュータ等のデバイスを用いて容易に利用することができる。また、新型コロナウイルス（COVID-19）の流行により、プライベートだけでなく、仕事においても対面におけるコミュニケーションが減少した。そのため、参加する場所の制約が少なく、同一の場所にいることなくコミュニケーションを行うことが可能であるビデオ会議システムの使用が劇的に増大した [KPA22, win20, Eva22]。

1.1 本研究で扱う多人数ビデオ会議の定義

ビデオ会議システムとは、2点以上の離れた場所にいる会議参加者間において、デジタル化されたビデオ画像および音声をやりとりするシステムである [WLB00]。ビデオ会議システムは仕事の場における「会議」の際に使用される。ここで「会議」を「3人以上の参加者が共通の場で対面し、情報の加工や伝達を行って、問題の解決を図ること」とする [高橋 87]。「会議」はその性質から、伝達会議、調整会議、決定会議、そして創造会議の4つに分類される [20008]。そのうち、本研究ではビデオ会議における課題が顕著に表れる創造会議に着目する。創造会議は新しいアイデアを創造するために行われる会議である。創造会議は、会議参加者が意見を活発に出すことが重要である。それに伴い、創造会議の会議参加者は発言数が増え、会議における話者交替が頻繁に起こる [井上 99]。以上のことから、本研究では創造会議に焦点を当てる。

また、多人数とは3人以上の複数人である。近年、ビデオ会議システムの発展に伴い、100人以上が参加する大規模なビデオ会議が実施されている。ビデオ会議はその性質から、1つのビデオ会議につき1つの話題に対する話しか行うことができない。そのため、大規模なビデオ会議において、会話を円滑に進めるため、ブレイクアウトセッションと呼ばれる3から10人程度の小規模なビデオ会議に分かれて会議を行う [WLV⁺21]。よって本研究が対象とする人数を3から10人とする。

以上のことから、本研究において、多人数ビデオ会議を「ビデオ会議システムを通して、3人から10人の参加者がそれぞれ別の場所からカメラを起動させて行う創造会議」と定義する。

1.2 多人数ビデオ会議における課題

創造会議において、話者交替が頻繁に起こる。しかし、ビデオ会議において話者交替は困難である。

対面会議においては、円滑な話者交替が可能である。これは会議の参加者が会話における各参加者の役割を示した参与役割を互いに把握しているためである [Gof76]。ここで参与役割とは、現在話している話し手、次に発話権を得るべき受け手、会話に参加しているが受け手とはならない傍参与者、および会話を聞いているだけの傍観者・盗み聞き者等が挙げられる。そして参与役割を把握する上で非言語情報が重要な働きを行う。非言語情報とは会話時における身振りや表情等の言語以外の情報である [Var86]。非言語情報の中で、特に視線配布（各参加者が誰を見ているかの情報）および配置は重要な役割を果たすことが知られている [Goo, Ken73]。しかし、ビデオ会議において、このような非言語情報は共有されない。そのため、参加者は互いの参与役割を把握することができず、話者交替が困難であると考えられる。またその結果、話者交替時の発話衝突や沈黙が多く発生すると考えられる。

ここで発話衝突とは、話し手が発話を終了した時、2人以上の聞き手が同時に話し始めてしまうことである。また、話し手が発話している状態において、聞き手が発話を開始することも発話衝突である。それに対して沈黙とは、話し手の発話が終了した時、聞き手が次に誰が発話するべきかを判断できず、誰も発話を行わない状態である。ビデオ会議において、非言語情報が共有されないことからこのような発話衝突や沈黙が多く発生する。

1.3 課題に対する先行研究

多人数ビデオ会議において、話者交替を円滑にすることを目的とした研究は多く存在する。

玉木ら [玉木 12] は、非言語情報から次に発話する可能性が最も高い参加者を推定し、強調表示することにより、円滑な話者交替を支援する手法を提案した。水野ら [水野 22] は頭部動作や視線等の非言語情報から受け手を予測している。また、著者の先行研究 [IKST21] において、視線から参与役割を推定し、ウィンドウサイズの大きさとして共有している。

これらの研究からビデオ会議においても、参与役割の把握が話者交替を促進することが明らかになった。しかし、これらの研究は発話欲求が高い参加者が複数いた場合、次話者となる受け手が複数人提示される。その結果、参加者の受け手の把握に差異が発生する可能性があり、話者衝突が発生すると考えられる。また、これらの研究は3人程度のビデオ会議を対象としており、それ以上の人数になった場合の参与役割の提示手法は検討されていない。以上のことから、本研究において発話欲求が高い参加者が複数人いる際においても話者交替を促進する手法を提案する。また、人数が増えた際においても参与役割の把握を促進させる提示手法を提案する。

1.4 目的とアプローチ

本研究の目的は多人数ビデオ会議において円滑な話者交替を促進することである。特に話者交替時に参与役割の把握を促すことにより、話者衝突および沈黙を減少させることを目的とする。そのために、ビデオ会議の参加者の参与役割を非言語情報から推定し、推定した参与役割をビデオ会議システムによって提示する手法を提案する。参与役割の推定について、参加者を話し手、受け手、傍参加者、傍観者の4つに推定する。提案手法において話者交替を促進するため、発話欲求が高い参加者のうち話し手の視線配布が行われている参加者を特に受け手として推定する。また、参与役割の提示について、社会学的知見であるF陣形を参考にビデオウィンドウの大きさおよび配置を用いて提示する。これによって、話者交替時に会議参加者が参与役割を把握することが促進され、円滑な話者交替が促進されることが考えられる。

1.5 貢献

本研究の貢献は以下の通りである。

- 多人数ビデオ会議における参加者の4つの参与役割（話し手、受け手、傍参加者、傍観者）を非言語情報から取得する手法を示した。
- 多人数ビデオ会議における参与役割をビデオウィンドウの配置および大きさを用いて提示する手法を示した。
- 提案手法を用いて実験を行い、本手法が会議に与える影響を調査し、話者交替を促進する可能性があることを示した。

1.6 本論文の構成

本論文の構成は以下の通りである。第1章においては、本研究の背景、本研究で扱うビデオ会議の定義、会議における課題、先行研究、目的とアプローチ、および貢献を示す。第2章においては、本研究で用いる社会学的知見、および本研究に関連する研究を述べ、本研究の位置付けを示す。第3章においては、本研究にて検討する手法の設計指針を示す。第4章においては、設計指針から実際に実装したシステムについて述べる。第5章においては、実装したシステムを評価するために行った実験について述べる。第6章において、実験の結果および考察について述べる。第7章においては、実験の結果から得た今後の課題について述べる。第8章においては、本研究の結論を述べる。

第2章 関連研究

本研究は、多人数ビデオ会議において話者交替を促進することを目的とし、そのために参与役割を利用する。本章では、多人数会議に関する社会学的知見、およびビデオ会議において非言語情報を利用した関連研究について説明した後、本研究の方針について述べる。

2.1 多人数ビデオ会議における社会学的知見

本研究では多人数ビデオ会議における話者交替を促進するために、対面での多人数会議に関する社会学的知見に着目した。本節ではまず会話に使用される非言語情報について述べ、次に話者交替について述べる。最後に、参与役割とF陣形について述べる。

2.1.1 会話に使用される非言語情報

非言語情報はコミュニケーションの中で重要な役割を果たしている [Var86]。ここで非言語情報とは、コミュニケーションを行う際の言語以外の情報のことを指す。言語によって伝えられるメッセージは全体の35%であり、残りの65%は非言語情報であると言われている [Bir10]。非言語情報は、対人距離、体の動き、表情、視線配布、接触、準言語、嗅覚作用、および人工物などが挙げられる [原岡 90]。本研究では特に視線配布および対人距離について取り上げる。これはビデオ会議において共有されていないことと話者交替において重要な役割を持つためである。Kendon [Ken67] は会話における視線の機能を3つに分類した。1つ目は注視の有無に基づいて発話を継続するか、終了したかを確認するモニタリング機能である。2つ目は聞き手の会話への興味を視線行動から察知し、発話内容を調整する機能である。3つ目は聞き手が会話への興味を表出させて、話し手へ伝える機能である。また、上記3つの他に、表情と組み合わせる感情を表現する、および対人関係の性質が表出するなど様々な機能もある [Var86]。

Hall [Hal92] は対人空間を密接距離、個体距離、社会距離、および公衆距離の4つの段階に分類し、その段階によって人間の振る舞いが区別されることを述べた。Hallが身体の距離に焦点を当てたことに対し、Kendon [Ken73, 坊農 08] は身体の空間配置行動に着目し、F陣形について述べた。F陣形とは会話構造を理解するための分析単位であり、2人以上の複数人が集まって会話する際の空間が維持される現象を説明している。

このように視線配布および対人距離は円滑な話者交替に必要な会話構造の把握に役立っている。この会話構造は参与構造と呼ばれている。本論文では [VSvdVN01] 項において参与構造を形成する参与役割とその把握を促す視線配布や身体配置 (F陣形) について詳細に述べる。

視線配布や身体配置が話者交替において重要な役割を担う一方で、それらはビデオ会議においてはほとんど機能していないと言える。よって視線配布および身体配置等の非言語情報を共有することによってビデオ会議において他の参加者の状態を把握するための判断材料が増え、より状況に適した会話を促進できるのではないかと考えられる。

2.1.2 話者交替について

話者交替とは、コミュニケーションを行う際に、現話者から次話者へ話者が変わることである。Sacks [SSJ78, 玉木 12] は話者交替のルールを下記のように定めた。

1. 「現話者による次話者選択」が話者交替時に行われた場合、選択された者が次話者となり話者交替が行われる。
2. 「現話者による次話者選択」が話者交替時に行われなかった場合、現話者以外の参加者は自己選択をし、最初に話し始めた者が次話者となる。
3. 「現話者による次話者選択」も自己選択も行われなかった場合、現話者が続けて話すことが可能である。しかし義務ではない。

ルール3が起これば現話者が続けて話した場合、次の話者交替時にルール1またはルール2が選ばれ、話者交替が起こるまでこれは繰り返される。

また話者交替の失敗には発話衝突、および沈黙が挙げられる。発話衝突とは複数の参加者が同時に発話を開始してしまうことであり、聞き手と聞き手間における発話衝突、または話し手と聞き手間における発話衝突が考えられる [小磯 00]。沈黙とは会話の参加者が誰も発話しない状態である。聞き手と聞き手間における発話衝突は話者交替のルール2において現話者以外の2人の参加者が同時に自己選択をすることにより発生すると考えられる。また、話し手と聞き手間における発話衝突は現話者が現話者以外の参加者の自己選択を認識できず、ルール2とルール3が同時に起こったと考えられる。沈黙について、ルール1において「現話者による次話者選択」に次話者の候補が気付かないため、もしくはルール2において自己選択を譲り合うなど考えられる。

ここで2.1.1項で述べた非言語情報の機能を考えると、話者交替の各ルールにおける判断に視線配布や身体配置が使われていると考えられる。すなわち、それらの共有されないビデオ会議においては、話者交替の各ルールにおける判断が困難になり、話者交替が失敗しやすいと考えられる。

2.1.3 参与役割とF陣形について

Goffman [Gof76] は会話の参加者が会話に参加する手続きを述べた (図 2.1)。

会話においては、現在話している話し手、次に発話権を得るべき受け手、会話に参加しているが受け手とはならない傍参与者、会話場への参与のみしている傍観者・盗み聞き者等、各参

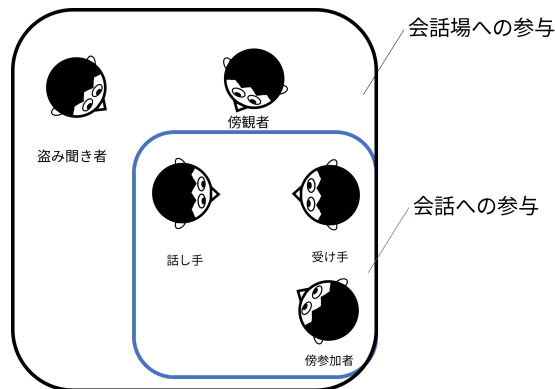


図 2.1: 会話における参与役割.

加者が参与の度合いに応じた役割（参与役割）を担う。そして話者交替が起きる際には話し手や受け手などの参与役割が動的に変化する。そして会話の参加者は互いの視線に基づいて自分の参与役割を理解していることが知られている [VSvdVN01, Ken67, 坊農 04, VSvdVN01, 佐藤 13]。話者交替における視線のやり取りを図 2.2 に示す。Kendon [Ken67] は話し手が発話の切れ目において受け手を注視し、受け手が注視を受け入れ話し手に注視を返すこと（相互注視）によって話者交替が成立することを示した。坊農ら [坊農 04] は、対面会話において会話へ参与する手続きを述べた（表 2.1）。

また、話し手の発話終了前の視線配布は発話権譲渡の意思を示し、聞き手が自らの参与役割を意識することを促すことを示した。会話の参与役割は話し手によって主体的に決められている。一方、聞き手の視線配布には、発話に対する興味、関心、および発話欲求が表出されており、話し手による次話者の選択に大きく影響している [VSvdVN01, 佐藤 13]。

また、対面会議において、各参加者の位置および方向といった身体配置は視線配布と同様に参与役割の把握に用いられる。F 陣形において、空間を複数人の会議参加者の下半身方向によって円陣の中央にできる O 型の空間（O 空間）、O 空間を構成するために参加者が身体を配置することによって作られる O 空間の外周的な空間（P 空間）、そして P 空間の外側の空間（R 空間）の 3 つに分ける。F 陣形における会話への参与の流れを図 2.3 に示す。

会話の途中から参与する人物はまず、R 空間に留まり、会話に参与するタイミングを見計らう（図 2.3 左）。その後、P 空間に身体を運び入れることで会話の参加者であることを他の会話の参加者に表明する（図 2.3 右）。ここで話者交替は P 空間で行われる。つまり、P 空間には話し手、受け手および傍参加者が存在し、O 空間に傍観者・盗み聞き者が存在する。会議参加者はこのようにして身体配置から参与役割を把握している。

以上より、視線配布および身体配置が共有されないビデオ会議において、参与役割を把握することは困難であり、話者交替が円滑に行われないと考えられる。よって本研究では、多人数ビデオ会議において、非言語情報から参与役割を推定し、推定した参与役割を提示する。これにより参与役割の把握を促し、話者交替が円滑に行われると考えた。

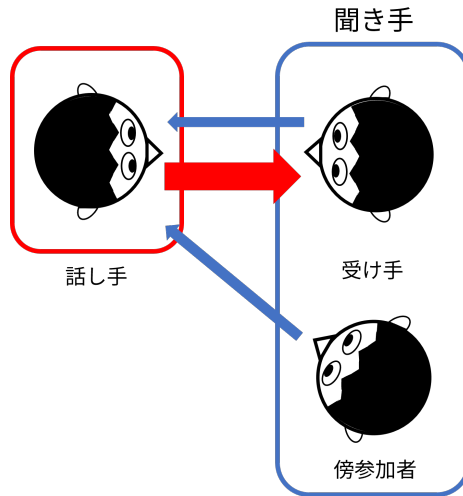


図 2.2: 話者交替時の視線のやり取り. 話し手と受け手は相互注視を行う. また, 傍参加者は話し手へ一方的に視線配布を行う.

表 2.1: 坊農ら [坊農 04] が述べた会話へ参与する手続き.

	行動	行動後の参与役割
(1)	意識的に会話場に近づき, ある一定時間留まる. 既存の参与者らに存在を意識され始める.	傍観者
(2)	既存の参与者らに確実に存在を意識される.	傍参加者
(3-a)	現話し手が, 自分に向けてデザインし産出した発話を受け取る.	受け手
(3-b)	意識的に既存の参与者に話しかける	話し手

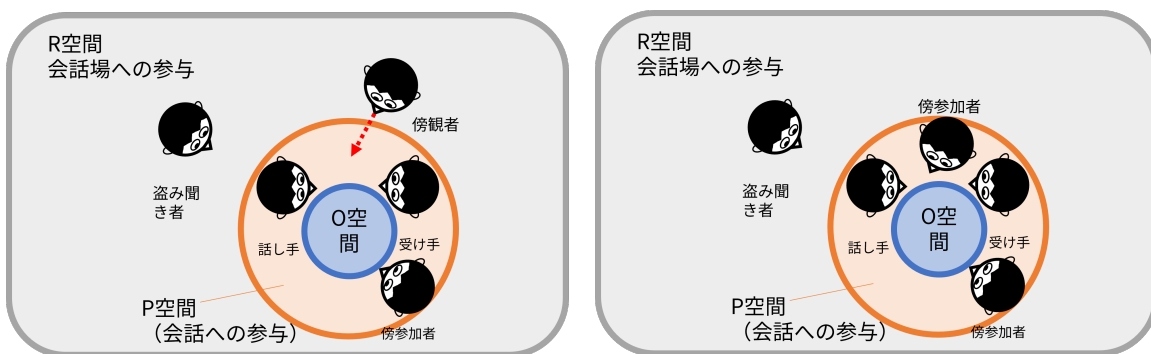


図 2.3: F 陣形における会話への参与の流れ. 会話へ途中から参与する人はまず R 空間にとどまる (左図). その後 P 空間に身体を運び入れる (右図).

2.2 ビデオ会議を支援する研究

ビデオ会議において非言語情報を用いて話者交替支援を行った研究は多く存在する。本節では現在普及しているビデオ会議システム、非言語情報を共有することで支援する研究、非言語情報を参与役割として提示する研究、および著者による先行研究について述べる。

2.2.1 現在普及しているビデオ会議システム

ビデオ会議システムの発展に伴い、大規模なイベントがビデオ会議システムを用いて開催されている。人数が大規模なビデオ会議において、会議参加者間の会話を促進するために、複数の小さな会議グループである分科会を形成する機能がビデオ会議システムに存在する。近年使用されるビデオ会議システムは、分科会を提供する手法において2種類に分類される [HAH22]。まず分科会を静的な枠組みとしてブレイクアウトルームを提供するビデオ会議システムである。これはあらかじめ準備されたそれぞれの部屋に参加者を振り分けることによって大規模なビデオ会議を支援する。これは他のグループの状況がわかりにくい一方、自分のグループへの会話に集中することができる。例えば Zoom [ZVC] や Microsoft Teams [Mic], Discord [Dis] などが挙げられる。もう一つは、一つの大きな空間が提供され、そこに自分のアイコンやアバター等を動かして動的に分科会を形成するビデオ会議システムである。例えば Gather [GP] や SpatialChat [Spa] 等が挙げられる。大規模なビデオ会議を行う際、このように分科会を複数作ることで会議を促進している。しかし、1つの分科会内で行われる会議は10人以下の小規模なビデオ会議であることが多く、1つの話題に対して会議を行う。また動的に分科会を形成システムは一つの大きな空間が提供されることにより、参加者個人のビデオウィンドウは小さくなる。そのため、先述したビデオ会議システムに比べて会話の負荷が高いと考えられる。また本研究で対象とするのは創造会議であり、会議へ集中することで会議が促進されると考えられる。よって本研究では1つの分科会で行われる10人以下の小規模なビデオ会議に着目する。

2.2.2 非言語情報を共有することで支援する研究

ビデオ会議において視線配布を伝達することにより会話の促進を目指した研究について述べる [Ver99, VWSC03, MOAY05, ACH96, VVdVV00, GM03, KKZ⁺17, 敷田 17, OMIM94]。

Vertegaal ら [Ver99, VWSC03] は、アイコンタクトを参加者に伝える多人数ビデオ会議システムを提案した。このシステムではディスプレイ内に表示された3次元CG空間内に各参加者のビデオウィンドウが立体的に配置され、ビデオウィンドウが回転することにより視線配布を提示する。また、Vertegaal ら [VVdVV00] は3人のグループに視線方向を共有しながら協調作業を行わせた。その結果、協調作業においても視線を共有することは話者交替を促進することを明らかにした。Mukawa ら [MOAY05] はビデオ会議においてカメラの向きを変え、視線を用いたシステムと視線を用いないシステムの比較を行い、ビデオ通話においても相互注視が有益であることを示した。視線を用いたシステムは対面会議と同様の視線行動が誘発

されることが明らかになった。Andersson ら [ACH96] はリモート会議参加者が存在する対面会議において、リモート会議参加者の存在感とアイコンタクトをローカル会議参加者へ提供した。Grayson ら [GM03] は1対1のビデオ通話において容易に相互注視を導入する方法を考案した。これは観測者に視線を向ける人が自分を見ているかを判断させることにより実現した。Kobayashi ら [KKZ⁺17] はグループ間のビデオ会議において、各ユーザの正しい視線方向を提示するシステムを提案した。画面埋め込みカメラや Kinect を用いて各参加者の適切な視点位置からの映像を生成し、他の参加者へ提示した。敷田ら [敷田 17] は参加者の1人だけ遠隔地から参加するビデオ会議において、リモート参加者の存在感を伝えるためにライトを視線情報に基づいて点灯させるウェアネス支援方式を提案した。Okada ら [OMIM94] は3人でのビデオ会議において等身大の相手画像と視線を一致させ、湾曲スクリーンに投影することにより、3人で机を囲んでいるようにビデオ会議を行った。佐藤ら [佐藤 11] はビデオ会議においてアイコンタクトを実現することによって円滑な会話を支援した。長谷川ら [長谷 14] は、発話意図表出に関連する無意識的身振りを伝達するテレプレゼンスロボットを提案した。この研究で用いられた無意識的身振りには頭部の回転も含まれ、テレプレゼンスロボットの顔部分が回転することにより、遠隔の参加者がどこを見ているかが伝達される。しかし、これらは2人での会話を想定したアイコンタクトであるものや、非言語情報を共有するためのシステムの導入に負荷が高いことが多い。そこで、本研究において複数人における会話の際に非言語情報を共有するシステムを提案する。

2.2.3 非言語情報を参与役割として表示する研究

玉木ら [玉木 12] は、Web 会議において、人が発話前に行う発話予備動作を検出し、次に発話する可能性が最も高い参加者を強調表示することにより、円滑な話者交替を支援する手法を提案した。ここで発話予備動作とは、発話前に体を前に乗り出す、および大きく頷くなど発話欲求を表出する動作のことである [Var86]。玉木らの提案システムにおいては、発話予備動作として手を上げる、手を顔の周りに近づける、および頷く動作が用いられており、視線方向の検出・提示は行われていない。また、発話欲求の高い参加者を強調するため、ビデオウィンドウをハイライトした。中野ら [中野 14] は多人数会話を行う会話エージェントの実装を行った。韻律情報と顔向き情報から受け手を推定するメカニズムを提案した。水野ら [水野 22] はビデオ会議において、頭部動作や視線等の非言語情報から受け手を予測している。山田ら [山田 22] は発話状態、および顔特徴から次話者である受け手を予測している。これらの研究について、推定する参加者自身の非言語情報から参与役割を推定している。そのため、受け手となりそうな参加者が同時に発生し、その結果発話衝突が発生すると考えられる。これに対し、本研究では話し手による視線配布から受け手を選択する。これにより発話欲求の高い参加者が複数人現れても、誰が話すべきが把握しやすいと考えられる。またこれらの研究は3人程度を想定されており、それを超える人数における参与役割を提示する手法については研究されていない。

2.2.4 著者による先行研究

著者の先行研究 [飯塚 22] において、多人数ビデオ会議において参与役割の把握を支援し、話者交替を促進するシステムを提案した。システムの実装にあたり、参与役割の推定および提示方法を検討した。まず、対面会議では参与役割の把握に視線配布が用いられるという知見に基づき、各参加者の視線情報を取得・提示するシステムを実装した。提示手法として、各参加者の視線情報より参与役割を推定しウィンドウサイズを変化させる手法（図 2.4 右）、および視線配布を直接矢印としてビデオウィンドウに重畳表示する手法（図 2.4 左）を実装した。

比較実験の結果、参与役割に応じてウィンドウサイズを変化させる手法において、参与役割の把握および話者交替が促進されることが示唆された。一方で、視線は不随意的な運動を伴うため、ウィンドウサイズやベクトル方向が頻繁に変化し会議への集中度が低下すること、参与役割の調整のための能動的な制御が困難であること、さらに参与役割の推定精度が不足している等の課題が示唆された。

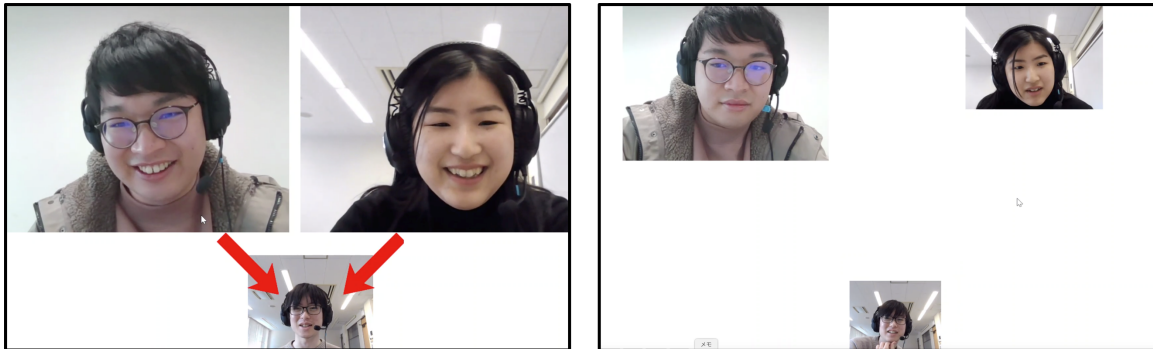


図 2.4: 先行研究における提案手法。ウィンドウサイズを変化させる手法（右図）。視線配布を直接矢印としてビデオウィンドウに重畳表示する手法（左図）。

2.3 本研究の位置付け

関連研究より、対面での多人数会議において、参加者の視線や配置等の非言語情報から参与役割を把握することによって円滑な話者交替を行うことができることがわかった。しかし、ビデオ会議において話者交替時に自分の参与役割が把握できず、発話衝突や沈黙が発生すると考えられる。これらの課題に対し、ビデオ会議において、非言語情報を共有する研究や参与役割を推定し提示することによって会話を促進する手法が研究されている。しかし、それらの研究は参与役割の推定精度が不十分であることや、受け手が複数人現れた際に円滑な話者交替を促進できないことが課題として考えられる。さらに、それらの研究は多人数の最小人数である 3 人における会議において検討されており、それ以上の人数に適した提示するための手法は検討されていない。そこで本研究では、多人数ビデオ会議において発話予備動作および視線配布を用いて話し手、受け手、傍参加者、および傍観者といった 4 つの参与役割を

推定する手法を提案する。また、推定した参与役割を F 陣形を参考に着想した画面表示（配置、ウィンドウサイズの変更）によって提示する手法を提案する。これらによって多人数ビデオ会議においても参与役割の把握を促し、話者交替を促進する。

第3章 設計指針

本研究は、多人数ビデオ会議における話者交替の促進を目的とする。そのために参与役割を推定し、提示する手法を提案する。関連研究において、参与役割の推定精度が不足していること、発話欲求が高い参加者が複数人いた場合について検討されていないこと、および3人より多い人数に適した提示手法は検討されていないことが課題として挙げられた。これらの課題に対し、本研究ではまず複数種類の非言語情報から参加者の参与役割が話し手、受け手、傍参加者、傍観者のうちどれに該当するかを推定する手法を提案する。また推定した参与役割を、F陣形を参考に着想した画面表示（配置およびウィンドウサイズの変更）によって提示する手法を提案する。本章では、提案手法の設計指針として、参与役割の推定手法および参与役割の提示手法について述べる。

3.1 参与役割の推定手法

提案手法において、参加者の非言語情報を利用して参与役割を推定する。参与役割 [VSvdVN01] として推定する役割は、話し手、受け手、傍参加者、および傍観者の4つとする。これらを選択した理由としてまず、話者交替に重要な役割であるため、話し手、受け手および傍参加者について、推定する。一方で傍観者および盗み聞き者は対面会話において、会話を聞いているが、積極的に会話に入らず、発話する意思のない参加者である。さらに盗み聞き者とは話し手に把握されていないが会話場へ参与している参加者である。これはビデオ会議において、カメラをオフにしたまま話を聞く参加者に当てはめることができると考えられる。これに対し本研究において、カメラをオンにして会議に積極的に参与する会議を対象としているため、盗み聞き者は推定する参与役割から外し、傍観者を推定する。

参加者の参与役割は非言語情報に応じて遷移する。推定する参与役割、推定する条件、推定の際に利用する非言語情報の関係を表3.1に示す。また、参与役割の遷移について図3.1に示す。まず、話し手の推定については、発話の有無（発話しているか否か）を用いる。参加者が発話をしていれば話し手であり、発話をしていなければその他の3つの参与役割のいずれかである。次に傍参加者、および傍観者の推定に関して、発話予備動作を用いる。発話予備動作とは発話の前に行う発話欲求の表出する動作 [Var86] である。過去の研究からビデオ会議において、頷き、開口、挙手、聞き手による話し手への視線配布 [玉木 12, 中野 14, 水野 22] などが発話予備動作として挙げられる。本研究ではこれに、直前の発話の有無を加えた5つの動作を組み合わせて利用する。5つの発話予備動作から発話したい度合いを表す発話欲求度を算出し、その大きさを発話欲求の高さを判断する。よって提案手法において、発話欲求

表 3.1: 提案手法における非言語情報と参与役割の関係

参与役割	推定する条件	利用する非言語情報
話し手	発話していること	発話の有無
受け手	発話欲求が高い, かつ話し手が視線配布している	発話予備動作, 話し手の視線配布
傍参加者	発話欲求が高い	発話予備動作
傍観者	発話欲求が低い	

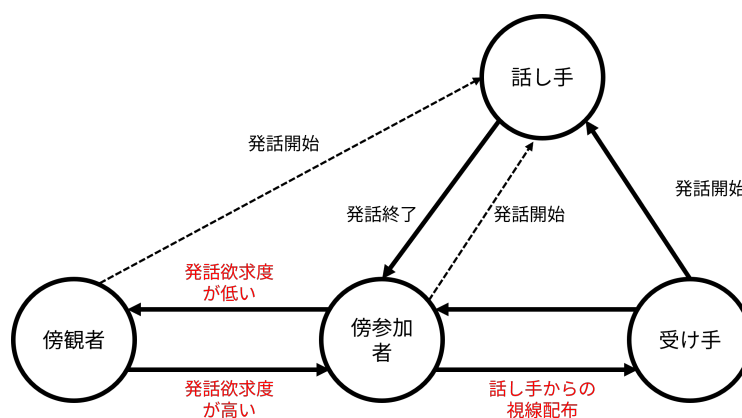


図 3.1: 参与役割の遷移. 実線の矢印は設計意図通りの遷移を示し, 破線の矢印は話者交替が失敗する可能性の高い遷移を表す.

が低い場合, 傍観者であると推定し, 発話欲求が高い場合, 傍参加者であると推定する. 傍参加者および受け手の推定に関して話し手の視線配布を使用する. 対面会議において, 話し手の視線配布によって主体的に受け手は選択される [VSvdVN01]. 対面会話において, 積極的に会話に参加し, 発話する意思があるもののうち, 話し手に選択され次の話者となるものが受け手であり, 話し手に選択されなかったものが傍参加者である. このことから提案手法において, 発話欲求が高い参加者である傍参加者のうち, 話し手が視線配布によって選択したものを受け手として推定する.

以上より, 複数の非言語情報を組み合わせて利用することにより, 先行研究 [IKST21] に比べ, 高い精度で参与役割が推定することが可能になると考えられる. また話し手が主体的に選択する受け手を加えた4つの参与役割を推定することにより, 発話欲求が高い参加者が複数人いた場合でも話者交替を促進することが可能になると考えられる.

3.2 参与役割の提示手法

推定した4つの参与役割を画面表示を用いて提示する。まず、ウィンドウサイズを変更することによって参与役割を提示する。話し手、受け手、傍参加者、そして傍観者の順でウィンドウサイズが大きく表示される。この順は、話者交替における重要度から決定した。先行研究 [IKST21] において、参与役割をビデオウィンドウサイズを用いて提示することによって話者交替が促進されることが明らかになっている。先行研究では主に話者交替が行われる3つの参与役割を提示している。しかし本研究においては傍観者を加えた4つの参与役割を提示するため、ウィンドウサイズを用いて提示することが十分であるかは明らかでない。そこで本研究では、ウィンドウサイズに加えてウィンドウの配置を用いて参与役割を提示する。配置によって参与役割を提示するために、社会学的知見であるF陣形(2.3) [Ken73]を参考に設計した。対面会議において、会議参加者は参与役割を把握するために身体配置を用いる。この時の参加者によって維持される空間を説明するためにF陣形が用いられる。F陣形において、空間を複数人の会議参加者の下半身方向によって円陣の中央にできるO型の空間(O空間)、O空間を構成するために参加者が身体を配置することによって作られるO空間の外円的な空間(P空間)、そしてP空間の外側の空間(R空間)の3つに分ける。会話の途中から参与する人物はまず、R空間に留まり、会話に参加するタイミングを見計らう。その後、P空間に身体を運び入れることで会話の参与者であることを他の会話の参与者に表明する。話者交替はP空間で行われ、P空間には話し手、受け手および傍参加者が存在し、O空間に傍観者が存在する。対面会議において会議参加者はこのような身体配置から参与役割を把握している。提案手法において、このP空間およびR空間を設定し、そこにビデオウィンドウを配置することによって参与役割を提示する。F陣形を上から見た際に円形となっていること、および空間を強調させるため、P空間およびR空間は円形にし、ビデオウィンドウをその中に円形に配置する。ウィンドウサイズおよび配置を組み合わせることで参与役割を提示することにより、人数が多いビデオ会議においても話者交替が行われる参加者が強調される。これにより、話者交替時に参与役割の把握を促進することが可能であり、話者交替が促進できると考えられる。

第4章 実装

前章において、多人数ビデオ会議において、参与役割を推定し、提示する手法の設計指針を示した。本章では、提案手法を実現するために実装したシステムについて述べる。まず、最初にシステム構成を述べ、非言語情報の検出、発話欲求度の算出、参与役割の推定、そして参与役割の表示について述べる。

4.1 ビデオ会議システムのシステム構成

実装したシステム構成を図4.1に示す。本研究では多人数ビデオ会議システムをWebアプリケーションとして実装した。これは導入する負荷が少ないビデオ会議システムを提供するためである。システムはReact, Typescript, HTML および CSS を用いて構成されている。ビデオ通話機能の実装について、WebRTCのプラットフォームであるSkyWay [Cor]を用いた。また、推定した参与役割の送受信、および視線配布の送受信についても、SkyWayを用いた。非言語情報を検出するために、様々な身体動作を取得可能なフレームワークであるMediaPipe [LTN⁺19]を用いた。まずMediaPipeのうち、Webカメラを用いて複数の身体動作をリアルタイムで同時に取得できるHolistic Solutionを用い、身体動作を取得した。MediaPipeはWebカメラから取得した画像から身体の特徴点における画像上の座標を取得する。Mediapipeから取得した身体動作はface landmarks および hand tracking である。face landmarks は機械学習を用いて3次元における468点の顔面ランドマークの座標を取得する。同様に、hand tracking は機械学習を用いて、3次元における21点の手のひらの座標を取得する。本システムでは、このように取得した身体動作から非言語情報を検出する。非言語情報の検出について4.2節に述べる。また、非言語情報のうち、直前の発話の有無を検出するため、Web Speech APIを用いた。検出した非言語情報から、参与役割を推定する。参与役割の推定に伴い、発話欲求の高低を発話欲求度として算出する。発話欲求度の算出について4.3節に述べる。参与役割の推定について4.4節に述べる。そして推定した参与役割をSkywayを用いて送受信を行う。受信した他の参加者の参与役割および推定した自分の参与役割を、提案手法を用いて参加者に提示する。参与役割の提示について4.5節に述べる。

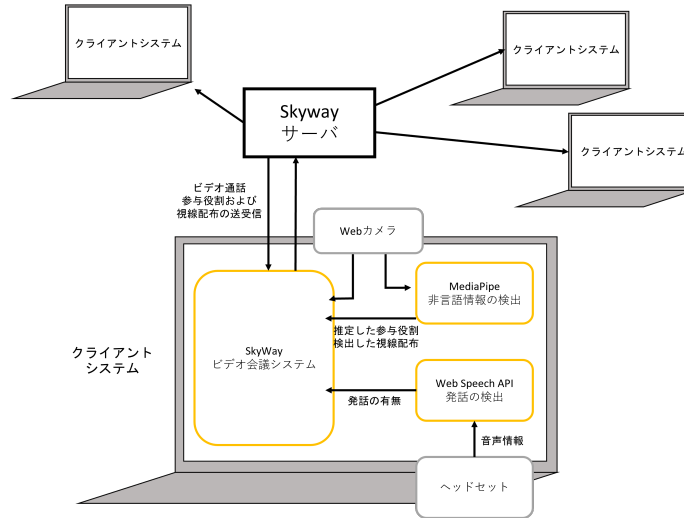


図 4.1: 本研究のシステム構成. 多人数ビデオ会議システムを Web アプリケーションとして実装した.

4.2 非言語情報の検出

提案システムにおいて, 参与役割を推定するために非言語情報を検出した. 非言語情報として, MediaPipe から取得した身体動作から手を挙げているか (以降, 挙手), 口を開けているか (以降, 開口), 頷きをしたか (以降, 頷き), および視線配布を検出する. また, Web Speech API を用いて発話をしたか (以降, 発話) を検出する. 本節において, 検出するそれぞれの非言語情報について述べる.

4.2.1 挙手の検出

挙手として手を開いて挙げている状態かどうかを検出した (図 4.2). 挙手を検出するために, MediaPipe の hand tracking を用いた. hand tracking で取得した 21 点の手のひらの座標のうち, 中指の先 h_1 および手のひらの底 h_2 の 2 点を用いた. まず, 手を開いて挙げている状態の中指の先 H_1 および手のひらの底 H_2 の距離 R_h を取得する. そして, Web カメラからリアルタイム取得した中指の先 h_1 および手のひらの底 h_2 の 2 点間の距離 r_h を取得する. この時, R_h と r_h が近い値である時, 挙手とした. 挙手を検出するための式を下記に示す.

$$R_h = |H_1 - H_2|$$

$$r_h = |h_1 - h_2|$$

$$f(x) = \begin{cases} \text{挙手} & (r_h \approx R_h) \\ \text{挙手でない} & (\textit{else}) \end{cases}$$

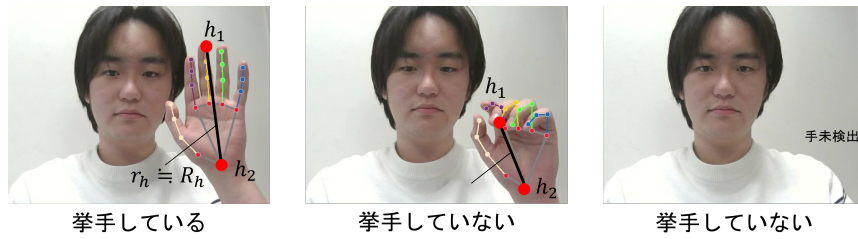


図 4.2: MediaPipe を用いた挙手の検出

検出した挙手は参与役割を推定するために算出する発話欲求のために使用する。

4.2.2 開口の検出

開口として口を開けている状態かどうかを検出した(図 4.3). 挙手を検出するために, MediaPipe の face mesh を用いた. face mesh で取得した 468 点の顔の座標のうち, 上唇の中心 m_1 および下唇の中心 m_2 の 2 点を用いた. Web カメラからリアルタイム取得した唇の中心 m_1 および下唇の中心 m_2 の 2 点間の距離 r_m を取得する. この時, r_m が 0 に近い値である時, 口を閉じている状態とし, そうでないときに開口とした. 開口を検出するための式を下記に示す.

$$r_m = |m_1 - m_2|$$

$$f(x) = \begin{cases} \text{開口でない} & (r_m \doteq 0) \\ \text{開口} & (\textit{else}) \end{cases}$$

検出した開口は参与役割を推定するために算出する発話欲求のために使用する。

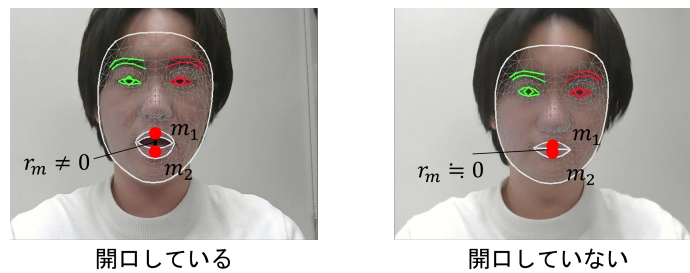


図 4.3: MediaPipe を用いた開口の検出

4.2.3 頷きの検出

頷きとして頷きが発生したかどうかを検出した(図 4.4). 頷きを検出するために, MediaPipe の face mesh を用いた. face mesh で取得した 468 点の顔の座標のうち, 正面を向いた際の顔の中心点として鼻の先である点 n を用いた. まず, 連続する 3 フレームにおける鼻の先の点の y 座標 n_{y1}, n_{y2}, n_{y3} を取得する. 次に連続した 2 フレーム間の顔の動きを取得する. 1-2 フレーム間の顔の動き r_{n1} を, および 2-3 フレーム間の顔の動き r_{n2} をとする. この時, r_{n1} が下向き, r_{n2} が上向き, かつ連続した 3 フレームにおける移動量 r_n が一定の値 k 以上の場合, 頷きが発生したとする. この時, k を Web カメラから取得した画像の縦の大きさを $height$ として,

$$k = height/10$$

とする. 頷きを検出するための式を下記に示す.

$$r_{n1} = n_{y2} - n_{y1}$$

$$r_{n2} = n_{y3} - n_{y2}$$

$$r_n = |r_{n1}| + |r_{n2}|$$

$$f(x) = \begin{cases} \text{頷きが発生} & (r_{n1} > 0, r_{n2} < 0, r_n > k) \\ \text{頷いていない} & (else) \end{cases}$$

検出した頷きは参与役割を推定するために算出する発話欲求のために使用する.

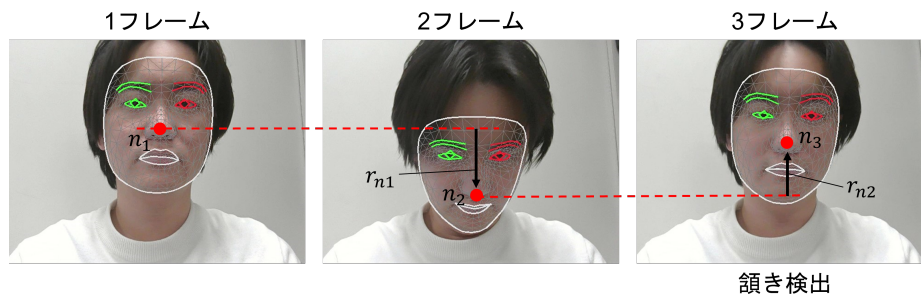


図 4.4: MediaPipe を用いた頷きの検出.

4.2.4 視線配布の検出

視線配布として顔がどこを向いているかを検出した(図 4.5). 先行研究 [IKST21] において, 視線配布としてディスプレイへの視線情報を用いることによって参与役割の推定にばらつきが発生し, 会議への集中が阻害される可能性があることが明らかになった. ここでディスプレイへの視線情報の代替手法として顔向きを取得する研究が多く存在する. そのため, 提案

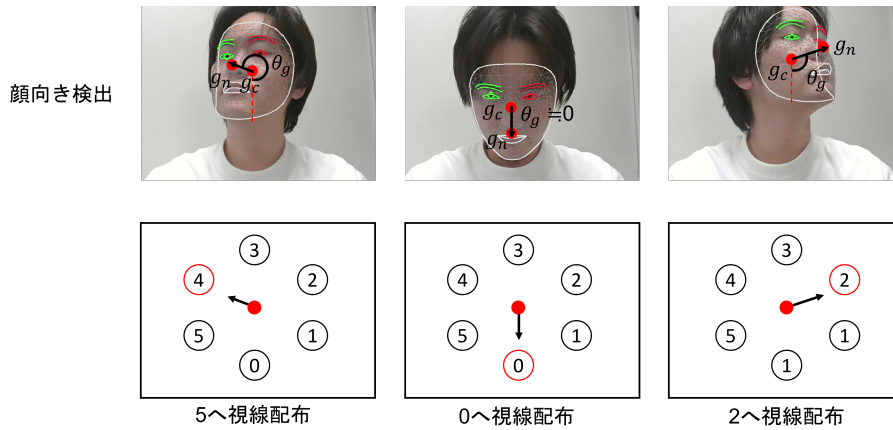


図 4.5: MediaPipe を用いた顔向き取得と視線配布の検出.

手法において、視線配布として顔向きを取得する。視線配布として顔向きを取得するために、MediaPipe の face mesh を用いた。まず face mesh からリアルタイムで取得した 468 点の顔の座標 $f_g(n)$ ($1 \leq n \leq 468$) から、顔の中心点 g_c を計算する。また、468 点の顔の座標のうち、正面を向いた際の顔の中心点として鼻の先である点 g_n を用いる。この時、顔の中心点 g_c を原点 $(0, 0)$ とした xy 座標における、鼻の先である点 g_n との角度を顔向き θ_g として取得する。顔向き θ_g を取得するための式を下記に示す。

$$g_c = \left(\sum_{1 \leq k \leq 468} f_g(k) \right) / 468$$

また、 g_c を原点とし、 $A(1, 0)$ とする。

$$\theta_g = \angle A g_c g_n$$

また、求めた顔向きから視線配布を検出する。視線配布としてディスプレイに円形に表示されたビデオウィンドウのうち、顔向き θ_g がどのビデオウィンドウに近い方向を向いているかを検出する。ビデオ会議への参加者の人数を n とする。 $x > 0$ における x 軸を含む範囲にビデオウィンドウが配置された参加者を 0 とし、反時計回りに参加者番号を割り振る。ある参加者における参加者番号を i ($0 \leq i < n$) とする。また、参加者番号 i である参加者への視線配布を g_i とする。この時、視線配布 g_i は下記の式で表される。

$$g_i = [\theta_g / (2\pi/n)]$$

検出した視線配布は、発話欲求度の算出および参与役割の推定における受け手の推定に使用する。

4.2.5 発話の検出

発話を取得するために Web Speech API を用いた。Web Speech API において、文字認識を行っている状態を発話中とし、文字認識を待機している状態を発話していない状態とした。検出した発話は発話欲求度の算出、および参与役割の推定における話し手の推定に使用する。

4.3 発話欲求度の算出

提案手法において参与役割である傍参加者の推定において、発話欲求の高さを発話欲求度として算出する。発話欲求度の算出には、前章で述べた非言語情報のうち、発話欲求の表出する動作である発話予備動作を利用する。提案手法では、発話予備動作として挙手、開口、頷き、話し手への視線配布、および直前の発話を用いる。それぞれの発話予備動作が検出されるごとに発話欲求度が増加する。発話予備動作の検出は1フレームごとに行われる。発話予備動作に対する発話欲求度の増減を計算する式は玉木らの研究 [玉木 13] を参考に設定した。

発話予備動作に対する発話欲求度を表 4.2 に示す。挙手について、手を挙げている時、発話欲求度を 5 加算する。開口について、口を開けている時、発話欲求度を 1.5 加算する。頷きについて、頷きが発生した時、発話欲求度を 15 加算する。話し手への視線配布について、話し手に視線配布を送っている時、発話欲求度を 0.5 加算する。最後に、直前の発話について、発話が終了した 1 フレーム後、発話欲求度を 100 とする。また、発話欲求度は 1 フレームごとに 0.5 ずつ減少する。発話欲求度は 0 から 100 の値をとる。あるフレーム x において、挙手を $f_h(x)$ 、開口を $f_m(x)$ 、頷きを $f_n(x)$ 、話し手への視線配布を $f_g(x)$ 、および直前の発話を $f_t(x)$ とした時、発話欲求度 $S(x)$ は下記の式で表される。

$$f_h(x) = \begin{cases} 5 & (\text{検出}) \\ 0 & (\text{未検出}) \end{cases}$$

$$f_m(x) = \begin{cases} 1.5 & (\text{検出}) \\ 0 & (\text{未検出}) \end{cases}$$

$$f_n(x) = \begin{cases} 15 & (\text{検出}) \\ 0 & (\text{未検出}) \end{cases}$$

$$f_g(x) = \begin{cases} 0.5 & (\text{検出}) \\ 0 & (\text{未検出}) \end{cases}$$

$$f_t(x) = \begin{cases} 1 & (\text{検出}) \\ 0 & (\text{未検出}) \end{cases}$$

$$S(x) = \begin{cases} 100 & (f_t(x) = 1) \\ S(x-1) + f_h(x) + f_m(x) + f_n(x) + f_g(x) - 0.5 & (f_t(x) = 0) \end{cases}$$

表 4.1: 発話予備動作と発話欲求度

発話予備動作	検出する動作	発話欲求度
挙手	手を挙げている	+5
開口	口を開けている	+1.5
頷き	頷きが発生	+15
話し手への視線配布	話し手へ視線配布を送っている	+0.5
直前の発話	直前に発話があった	=100

4.4 参与役割の推定

提案手法において、参加者を話し手、受け手、傍参加者、および傍観者の4つの参与役割に推定する。本節では、3.1節で述べた設計指針に則り、4.2節で述べた非言語情報および4.3節で述べた発話欲求度を用いて参与役割を推定する手法を示す。

話し手の推定

非言語情報のうち、発話の検出を用いる。発話を検出し続けている状態の時、参加者の参与役割を話し手とする。発話を検出している時、発話欲求度は算出されない。発話が終了した時、傍参加者に遷移し、発話欲求度は100から減少していく。

受け手の推定

発話欲求度、および非言語情報のうち話し手の視線配布の検出を用いる。発話欲求度が50以上であること、かつ話し手となっている参加者から視線配布が行われていた場合、参加者の参与役割を受け手とする。話し手からの視線配布が行われなくなった場合、傍参加者に遷移する。発話欲求度が50を下回った場合、傍観者に遷移する。

傍参加者の推定

発話欲求度、および非言語情報のうち話し手の視線配布の検出を用いる。発話欲求度が50以上であること、かつ話し手となっている参加者から視線配布が行われていない場合、参加者の参与役割を傍参加者とする。話し手の視線配布が行われた場合、受け手に遷移する。発話欲求度が50を下回った場合、傍観者に遷移する。また発話を行った場合、話し手に遷移する。

傍観者の推定

発話欲求度を用いる。発話欲求度が50を下回っていた場合、参加者の参与役割を傍観者とする。発話欲求度が50以上となった場合、傍参加者に遷移する。また発話を行った場合、話し手に遷移する。

4.5 参与役割の提示

提案手法において推定した4つの参与役割を画面表示を用いて提示する。画面表示はビデオウィンドウサイズ（以降、ウィンドウサイズ）および配置を用いる。実装したシステムを図4.6に示す。本節において、ウィンドウサイズを用いた提示、および配置を用いた提示について述べる。



図 4.6: 実装したシステム。画面表示を用いて参与役割を提示する。

4.5.1 ウィンドウサイズを用いた提示

ウィンドウサイズの大きさを用いて参与役割を提示する。ウィンドウサイズは話し手、受け手、傍参加者、そして傍観者の順で大きく表示される。表 4.2 にウィンドウサイズの大きさを述べる。サイズはビデオ会議システムにおけるディスプレイサイズの横幅を 100%とした時の大きさである。ウィンドウサイズの大きさは4つの大きさの違いが明確に認識できるように著者が設定した。また、傍観者が目立たず、話し手へ視線配布が誘導できるように設定した。

表 4.2: ウィンドウサイズの大きさ.

参与役割	ウィンドウサイズ
話し手	16%
受け手	12%
傍参加者	8%
傍観者	4%

4.5.2 配置を用いた提示

ウィンドウサイズに加え、配置を用いて参与役割を提示する。ウィンドウサイズと配置を組み合わせることで、より参与役割の把握を促進できると考えられる。配置はF陣形を参考に設計した。具体的には、R空間およびP空間を分けて提示する。R空間には傍参加者が存在し、会議への参与が小さい参加者を提示する。また、P空間は話者交替が行われる話し手、受け手、傍参加者が存在し、会話への参与が大きい参加者を提示する。ディスプレイの中心を円の中心としてビデオウィンドウを円形に配置し、その半径を変更することでP空間およびR空間を提示する。表 4.3 に円形に配置する際の半径を述べる。半径はビデオ会議システムにおけるディスプレイサイズの縦幅および横幅をそれぞれ 100%とした時の長さとした。半径はP空間およびR空間の違いが明確に認識できるように著者が設定した。また、話者交替が行われるP空間を強調するため、P空間を黄色の円を表示する。P空間は半径が 25%とした。P空間の色は注意を引くために、黄色に設計した。

表 4.3: ウィンドウを円形に配置する際の半径と参与役割.

参与役割	半径
話し手, 受け手, 傍参加者	17.5%
傍観者	45%

第5章 実験

前章の実装したシステムを用いて、提案手法が多人数ビデオ会議において話者交替を促進するのかを明らかにするために実験を行った。本章では、実験設計、実験手順、仮説と評価、および結果を述べる。

5.1 実験設計

本実験において、6人の実験参加者がビデオ会議システムを用いて創造会議を行う実験設計とした。会議の種類は話者交替が頻繁に行われることから創造会議とした。会議における参加者の人数について、Davisは5人以上から議論への活動が増加することを報告した [Dav73]。Gorseらは会議の人数は5人を下回るべきではないことや6人が適しているといった知見があることを述べた [GS07]。また、Kooloosらは議論のバランスおよび個人の能力開発の促進等の点から5、6人が適した人数であることを述べた [KKV+11]。本実験においては、創造会議を対象とし活発に議論を行い参加者が多くの意見を出すことが重要であるため、会議へ参加する人数を6人とした。

5.1.1 実験条件

本実験において、提案手法が話者交替を促進するのかを明らかにするために、3つの条件を用いて比較を行った。提案手法を実装したシステムを用いた参与役割提示（サイズ+配置）条件、提案手法を一部実装したシステムを用いた参与役割提示（サイズ）条件、および既存のビデオ会議システムにおける手法を用いた話し手強調提示条件を実験条件（図 5.1）とした。各条件における詳細を下記に記す。

参与役割提示（サイズ+配置）条件

参与役割提示（サイズ+配置）条件では、前章にて述べた提案手法を全て実装したシステムを用いて会議を行う。参与役割提示（サイズ+配置）条件で用いたビデオ会議システムの外観を図 5.1 上に記す。参与役割提示（サイズ+配置）条件において、システムは非言語情報を元に参加者の参与役割（4種類）を推定する。推定した参与役割をビデオウィンドウサイズ、および配置を用いて提示する。また話者交替が行われる P 空間を強調するため、背景に黄色の円を表示する。

参与役割提示（サイズ）条件

参与役割提示（サイズ）条件では、前章にて述べた提案手法を一部実装したシステムを用いて会議を行う。参与役割提示（サイズ）条件を用いて行ったビデオ会議システムを図 5.1 下左に記す。参与役割提示（サイズ）条件において、システムは非言語情報を元に参加者の参与役割（4 種類）を推定する。推定した参与役割をビデオウィンドウサイズを用いて提示する。各参加者のビデオウィンドウは同一円周上に配置される。ビデオウィンドウが配置される円の半径は、ビデオ会議システムにおけるディスプレイサイズの縦幅および横幅をそれぞれ 100%とした時の 35%の長さとした。

話し手強調提示条件

話し手強調提示条件では、既存のビデオ会議システムにおける手法を用いて会議を行う。話し手条件を用いて行ったビデオ会議システムを図 5.1 下右に記す。話し手条件において、システムは非言語情報を元に参加者の参与役割（4 種類）を推定する。しかし、推定した参与役割のうち、話し手がどうかのみビデオウィンドウサイズを用いて提示する。ビデオウィンドウサイズは、ビデオ会議システムにおけるディスプレイサイズの横幅を 100%とした時、話し手が 16%であり、その他の参与役割を 8%とした。ビデオウィンドウの配置については参与役割提示（サイズ）と同様である。

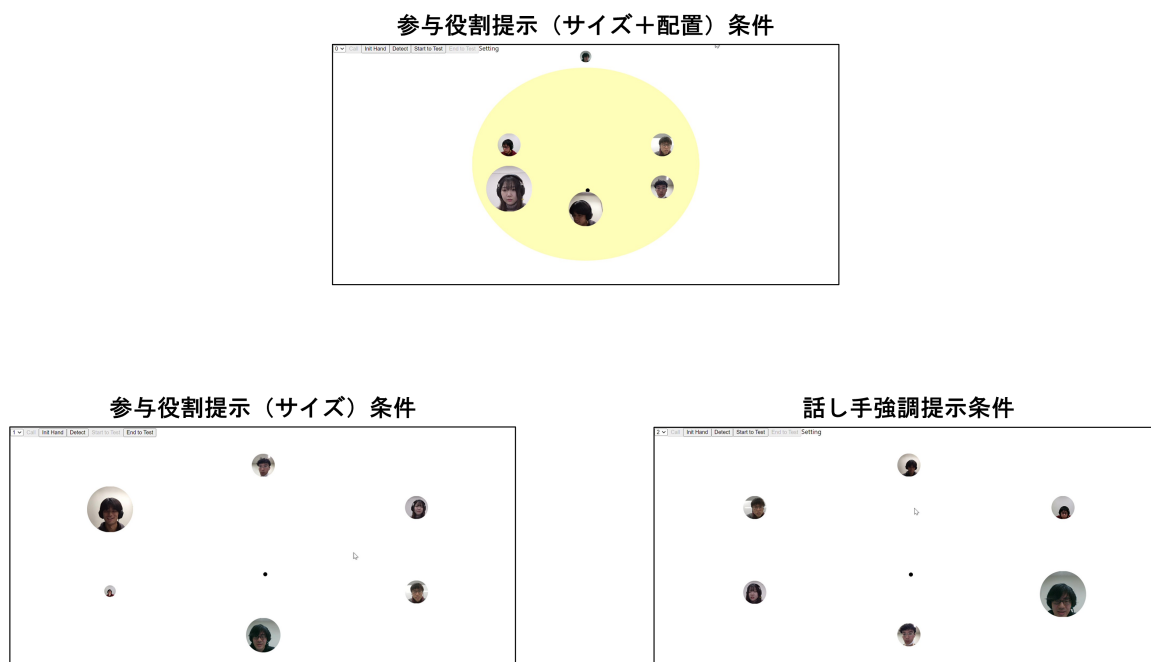


図 5.1: 実験条件. 上：参与役割提示（サイズ+配置）条件. 下左：参与役割提示（サイズ）. 下右：話し手強調提示条件.

5.1.2 実験タスク

本実験では、実験タスクである創造会議として、Guilford's Alternative Uses Task（以降、AUT）[MP59,AMH⁺,Cre22]を行った。AUTは対象者の創造性を測定するためのテストであり、提示された特定の物（例えば鍋など）の新しい用途を参加者で議論する。AUTにおいて、流暢性、独創性、柔軟性、綿密性の4つの側面から創造性を評価する。本実験では多人数においてAUTに加え、その会議の会話分析を行い、話者交替が促進されたかを評価する。

AUTに基づき、参加者は題材に対する新しい用途をできるだけ多く考えた。本実験において、題材を“鍋”、“靴下”、“レンガ”に設定した。これは、関連研究[櫻井15,Cre22,山岡17,中原19,NYS⁺14]から以下の理由を基準に選択した。

- 題材について全く知らないことがなく、詳しすぎることはない
- 題材について新しい用途を考える課題をしたことがない
- 題材についてある程度新しい用途を考えることができる
- 各々の題材が別の題材の回答に応用できることがほとんどない

実験は参加者内配置において行った。1時間未満の会議において、時間経過とともに発話回数は減少しない[玉木13]とされていることから、1つの条件につき、会議を7分間行い、合計21分間会議を行った。

タスク実施時は、参加者にシステムが提示する情報を参考に会議を進めるよう指示した。また、議論に当たって参加者の中で司会を設定せずに会議を行うよう指示した。これは、司会となる参加者が他の参加者に会話を振ることにより、システムを用いた話者交替が減少することを避けるためである。

各参加者はそれぞれ与えられた部屋に分かれて会議を行った。参加者はそれぞれラップトップPCおよびヘッドセットを用いて会議を行った。

5.1.3 実験参加者

実験参加者は12名（男性11名、女性1名、平均23歳）であった。参加者は6名1組の2グループに分けられた。グループ内における参加者同士は互いに知り合いとした。これは初対面であることに起因する会話への負荷をなくすためである。各グループにおける条件および題材の順序を表5.1に示す。順序による影響を考慮し、グループごとに順序を変更した。

表 5.1: 各グループにおける条件および題材.

	1 回目	2 回目	3 回目
A	参与役割提示 (サイズ+配置) 条件 題材: レンガ	参与役割提示 (サイズ) 条件 題材: 靴下	話し手強調提示条件 題材: 鍋
B	参与役割提示 (サイズ) 条件 題材: 鍋	話し手提示提示 題材: レンガ	参与役割提示 (サイズ+配置) 条件 題材: 靴下

5.2 実験手順

実験全体の手順を図 5.2 に示す.

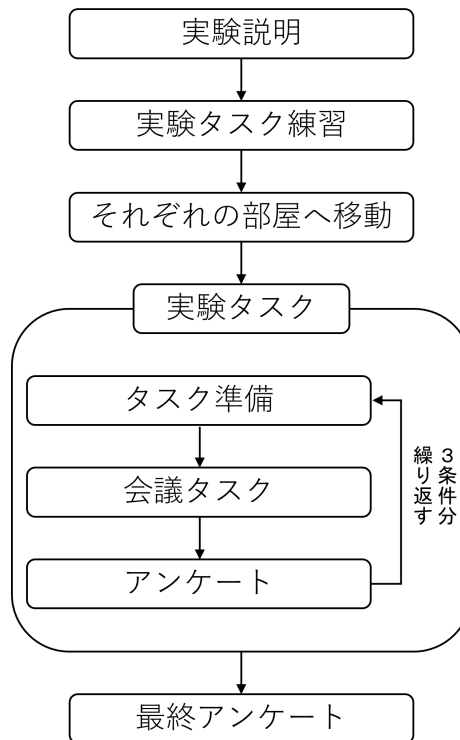


図 5.2: 実験手順.

まず実験実施者は参加者に対して実験タスクについての説明を行った。その後、システムについての説明を行った。本実験において、参加者はそれぞれの与えられた部屋に分かれてビデオ会議を行う。よってビデオ会議を行うための準備を参加者自身でやる必要があった。参加者は説明に従って一度ビデオ通話を接続し、非言語情報の取得までの操作を行った。システムについての説明が終了した後、実験タスクの練習を行った。これは、タスク実施1回目とそれ以降では、タスクや他の参加者への慣れの差により振る舞いに変化する可能性があると考えたためである。実験タスクの練習では、参加者は題材を‘輪ゴム’としたAUTを対面環境で行った。この際、会議時間は2分間とした。実験タスクの練習が終了した後、参加者はそれぞれ与えられた部屋に分かれる。全員が与えられた部屋に着き次第、1度目の実験タスクのための準備を行う。参加者全員のタスク準備が終了した後、実施者の合図に従ってタスクである会議を始める。7分間の会議を行い、実施者の合図に従って会議を終了する。その後、それぞれの条件についてのアンケート（表5.2）を行う。アンケートの回答後、次の条件でタスク準備を行う。このタスク準備からアンケートまでを実験タスクとし、これを3回繰り返す。全ての条件でのビデオ会議を終え、3回目のアンケートの回答後、最終アンケート（表5.3）に回答する。最終アンケートの回答後、実験が終了する。1回の実験につき所要時間は約90分であった。

表 5.2: 各条件実施後に参加者に回答させたアンケートを示す。AUTにおける創造性を評価(a)するための4つの項目を10段階リッカート尺度を用いた回答欄にて回答させた。また直前まで行っていた会議における参与役割の把握(b)および会議(c)についての印象評価を5段階リッカート尺度を用いた回答欄にて回答させた。

番号	項目
a-1	アイデアの数はたくさん考えられたか
a-2	独創的なアイデアを考えることができたか
a-3	幅広い用途のアイデアを考えることができたか
a-4	詳細なアイデアを考えることができたか
b-1	今自分が会話に参加していること/していないことを意識することができた
b-2	自分の話を聞いている参加者がわかった
b-3	自分がちゃんと聞いていることが相手に伝わっているように感じた
b-4	自分が話し終えるとき、次に話たさそうな人がわかった
b-5	自分が話し終えるとき、話したくない人がわかった
b-6	他の人が話し終えるとき、自分が話すべきかがわかった
c-1	議論に集中できた
c-2	議論にストレスを感じた
c-3	議論に疲労を感じた
c-4	話者交替にストレスを感じた

表 5.3: 全ての条件が終了した後に参加者に回答させたアンケートを示す。ここでは提案手法に対する印象を5段階で回答させた。またシステム、および実験に関する印象を記述式に回答させた。

番号	項目
d-1	条件0, 1において, 自分が話したい時にシステムはそれを示せたか
d-2	条件0, 1において, 自分が話したくない時にシステムはそれを示せたか
e-1	システムが会話に与えた影響について何か感じたことがあれば記入してください
e-2	その他, 今回のタスクの中で気づいたことや問題点があれば記入してください
e-3	話者交替(話者衝突, 沈黙など)について何か感じたことがあれば記入してください

5.3 仮説と評価項目

本実験では、評価のために各参加者の会議時間における発話内容、参与役割、および発話欲求度と発話予備動作を記録した。また、会議の音声およびビデオ画面を記録した。本実験ではこれらを用いて会話分析を行う。評価項目として、話者交替数、発話衝突の回数および確率、沈黙時間、参加者の発話数、参与役割の遷移、発話欲求度の平均、および発話予備動作を用いる。また、参加者が回答したアンケートを用いて、会議および参与役割について印象評価を行う。

本実験では、以上の評価項目を用いて次の3つの仮説の検証を行う。

5.3.1 仮説1：参与役割提示条件では、話し手強調提示条件と比較して、参与役割の把握が促進される。

参与役割提示条件(サイズ+配置, サイズ)においては、推定した4種類の参与役割をウィンドウサイズおよび配置を用いて提示する。これにより、話し手か否かのみを提示する話し手強調提示条件と比較して、参与役割の把握が促進されることが考えられる。また参与役割提示条件(サイズ+配置)では、ウィンドウサイズに加えて配置を用いて参与役割を提示することによって参与役割の把握が促進されることが考えられる。

この仮説を検証するために会話分析を行う。会話分析では、参与役割の把握を促進したかを評価するために参与役割の遷移を評価する。具体的には、話し手の前に受け手であったか、受け手の後に話し手であったか、傍参加者の後に傍観者であったか、傍参加者の後に話し手であったか、および傍観者の後に話し手であったかをその回数から評価する。これは、参与役割が適切に把握される場合、受け手が次の話し手になる割合が高く、そこから傍参加者、傍観者の順に割合が減少すると考えたためである。また、参与役割の把握を参加者の行動から評価するため、発話欲求度の平均、および発話予備動作について評価する。発話予備動作として、話し手への視線配布の時間、挙手した回数および時間、頷き回数、開口した回数を評価する。また、アンケートにより、参与役割の把握に対する印象を評価する。

5.3.2 仮説2：参与役割提示条件では話し手強調提示条件と比較して、円滑な話者交替が促進される。

参与役割提示条件（サイズ+配置，サイズ）において参与役割の把握が促進される場合（仮説1），話し手強調提示条件と比較して，円滑な話者交替が促進されると考えられる。

この仮説を検証するために会話分析を行う。話者交替を促進したか評価するために発話数および話者交替数を評価する。また話者交替の失敗が減少したかを評価するため，発話衝突の回数および確率，沈黙時間を評価する。また，アンケートにより，話者交替への印象を評価する。

5.3.3 仮説3：参与役割提示条件では話し手強調提示条件と比較して，会議の質が向上する。

参与役割提示条件において，円滑な話者交替が促進される場合（仮説2），議論することの負荷が減少し，会議の質が向上すると考えられる。

会議の質が向上したかを評価するため，アンケートにより会議への印象を評価する。それに加え，AUTによる創造性への影響を評価する。

第6章 実験結果および考察

本章では会話分析およびアンケートによる印象評価の結果について述べる。

6.1 会話分析

本項では会話分析として、参加者の発話数、話者交替数、発話衝突の回数および確率、沈黙時間、参与役割の遷移、発話欲求度の平均、および発話予備動作を評価した結果を述べる。

6.1.1 発話数

各条件における参加者の発話数の平均を図 6.1 に示す。グラフのエラーバーは標準偏差を示す。発話数は各参加者がシステムを用いた7分間の会議内において、発話を行った回数を表す。各参加者の発話数に対して Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。次に一元配置分散分析を行ったところ有意差が示された ($p = 0.013 < 0.050$)。Bonferroni-Holm の方法により有意水準を補正し、多重比較を行ったところ、各条件間に有意差は示されなかった。

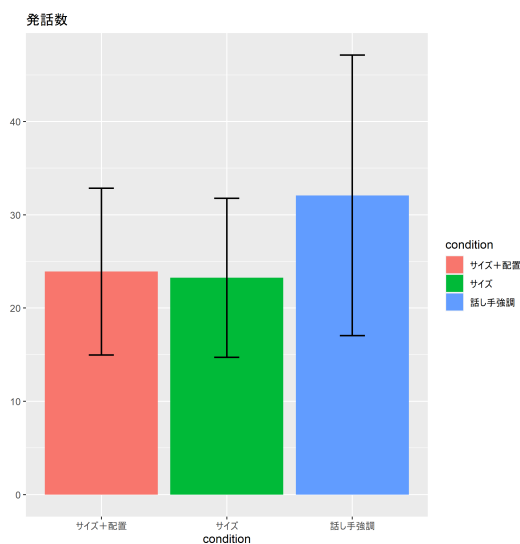


図 6.1: 各条件における参加者の発話数の平均。

6.1.2 話者交替数

各条件における参加者の話者交替数の平均を図 6.2 に示す。グラフのエラーバーは標準偏差を示す。話者交替数は各参加者がシステムを用いた 7 分間の会議内において、自分の発話の後に他の人が発話した回数を表す。各参加者の発話数に対して Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。次に一元配置分散分析を行ったところ有意差が示された ($p = 0.0082 < 0.010$)。Bonferroni-Holm の方法により有意水準を補正し、多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.046 < 0.050$)。また、参与役割提示 (サイズ) 条件-話し手強調提示条件間において有意差が示された ($p = 0.032 < 0.050$)。このことから参与役割提示条件に対して、話し手強調提示条件は話者交替数が有意に多いことが示された。

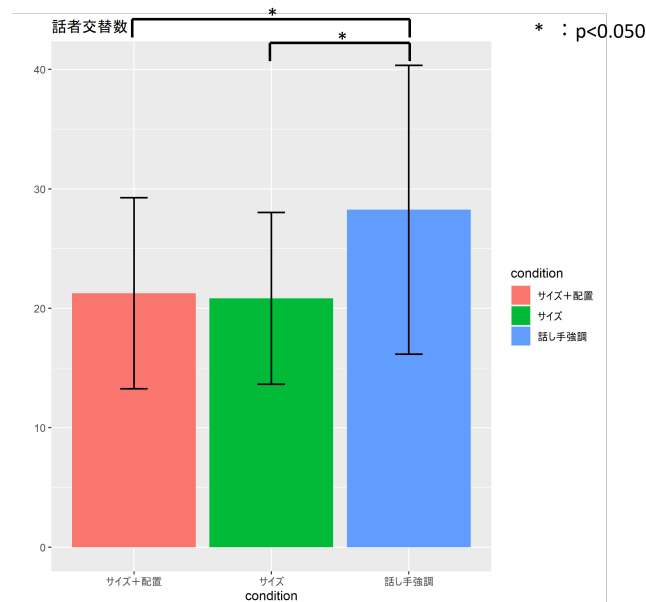


図 6.2: 各条件における参加者の話者交替数の平均。

6.1.3 発話衝突の回数および確率

各条件における参加者の発話衝突の回数の平均を図 6.3 に示す。各条件における参加者の発話衝突の確率を図 6.4 に示す。エラーバーは標準偏差を示す。本研究では、発話衝突について、自分が発話を開始した際に他の人が既に発話していた数 (図 6.3A) および自分が発話している際に他の人が発話を開始した数 (図 6.3B) に分類した。また、各参加者の 2 種類の発話衝突の回数を足し、発話数で割った数を発話衝突の確率 (図 6.4) とした。自分が発話を開始した際に他の人が既に発話していた際の発話衝突の回数 (A) に対して Shapiro-Wilk 検定

を行った結果、非正規性が示された。そこで、変数変換を行うことにより非正規性の変数において一元配置分散分析を行うことが可能となる Aligned Rank Transform [WFGH11, HBT90] (以降, ART) を行い、一元配置分散分析を行った。一元配置分散分析の結果、有意差が示された ($p = 0.011 < 0.050$)。多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.010 < 0.050$)。よって、A の発話衝突の回数は参与役割提示 (サイズ+配置) 条件は話し手強調提示条件に比べて有意に少ないことが示された。自分が発話している際に他の人が発話を開始した際の発話衝突の回数 (B) に対して Shapiro-Wilk 検定を行った結果、非正規性が示された。そこで、ART を行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.0046 < 0.010$)。多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.0048 < 0.010$)。また、参与役割提示 (サイズ) 条件-話し手強調提示条件間において有意差が示された ($p = 0.037 < 0.050$)。よって、B の発話衝突の回数は参与役割提示条件は話し手強調提示条件に比べて有意に少ないことが示された。発話衝突の確率に対して Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。次に一元配置分散分析を行ったところ有意差が示された ($p = 0.012 < 0.050$)。Bonferroni-Holm の方法により有意水準を補正し、多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.0091 < 0.010$)。このことから参与役割提示 (サイズ+配置) 条件は、話し手強調条件に比べて発話衝突の確率は有意に少ないことが示された。

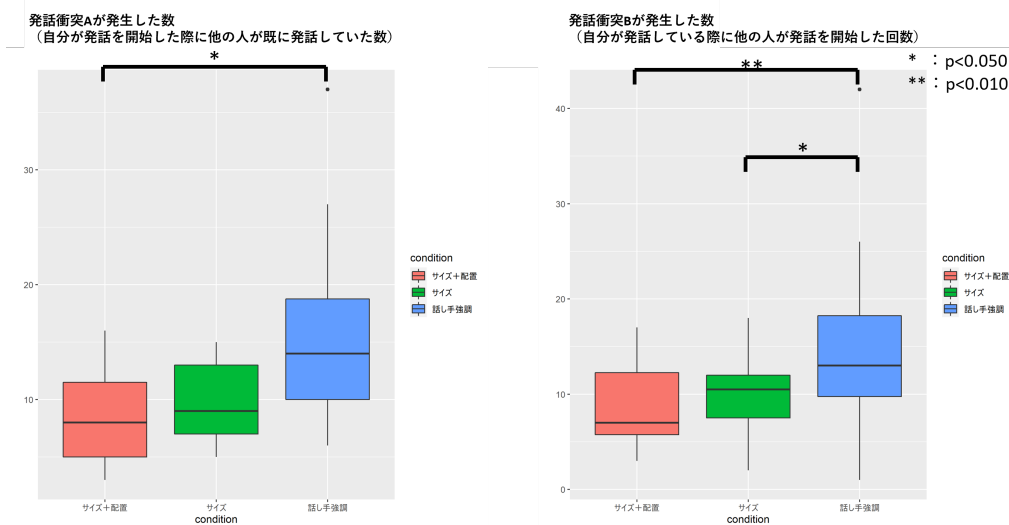


図 6.3: 各条件における参加者の発話衝突の回数の平均。A：自分が発話を開始した際に他の人が既に発話していた数。B：自分が発話している際に他の人が発話を開始した数。

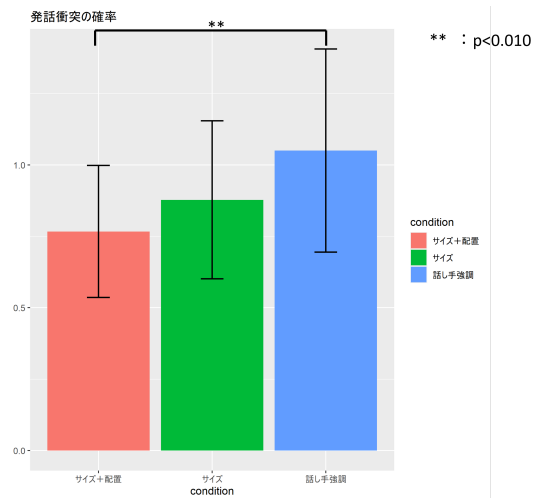


図 6.4: 各条件における参加者の発話衝突の確率.

6.1.4 沈黙時間

各条件における各グループの沈黙時間の平均を図 6.5 に示す. グラフのエラーバーは標準偏差を示す. 沈黙は各グループがシステムを用いた 7 分間の会議内において, 1 人も参加者の発話が無かった時間を表す. 各グループの沈黙に対して Shapiro-Wilk 検定を行った結果, 正規性が示された. そのため, Bartlett 検定を行った結果, 等分散性が示された. 一元配置分散分析を行ったところ有意差が示されなかった ($p = 0.7472$).

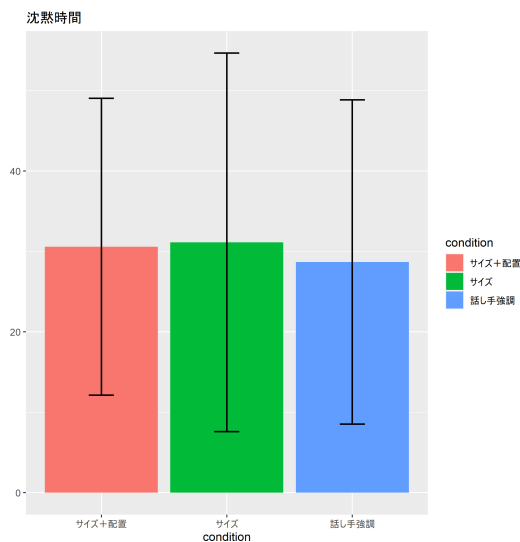


図 6.5: 各条件における各グループの沈黙時間.

6.1.5 参与役割の遷移

- 話し手の前に受け手であったか

話し手の前に受け手であったかの評価を図 6.6 に示す。具体的には、話し手となる前の 5 秒以内に受け手となった回数を評価する。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意傾向が示された ($p = 0.097 < 0.10$)。多重比較を行ったところ、有意差は示されなかった。

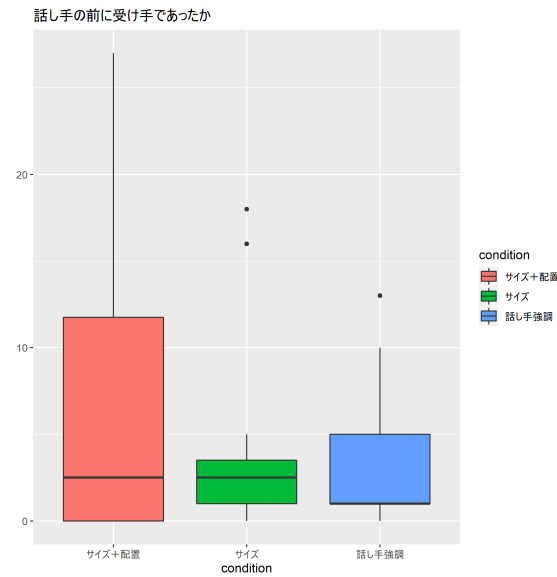


図 6.6: 話し手の前に受け手であったかの評価

- 受け手の後に話し手であったか

受け手の後に話し手であったかの評価を図 6.7 に示す。具体的には、受け手となった後 5 秒以内に話し手になった回数を評価する。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意傾向が示された ($p = 0.098 < 0.10$)。多重比較を行ったところ、有意差は示されなかった。

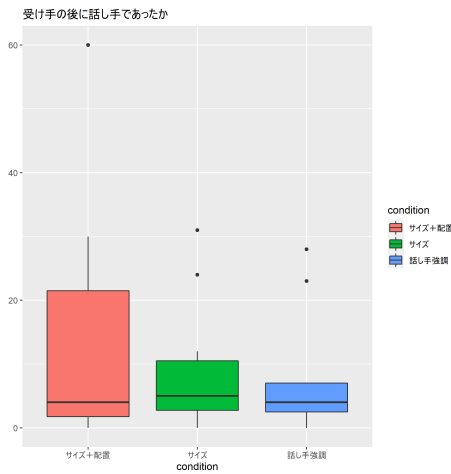


図 6.7: 受け手の後に話し手であったかの評価

- 傍参加者の後に傍観者であったか

傍参加者の後に傍観者であったかの評価を図 6.8 に示す。具体的には、傍観者から傍参加者になり、その後話し手になることなく傍観者になった回数を評価する。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意差は示されなかった ($p = 0.11$)。

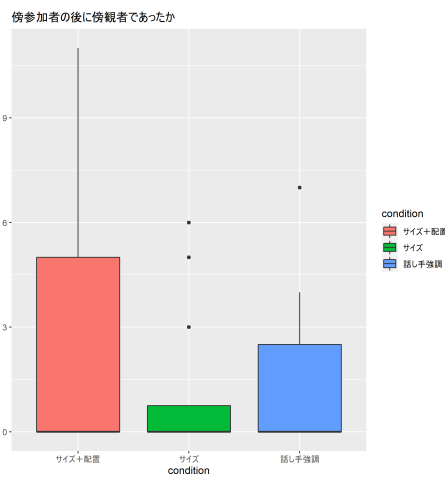


図 6.8: 傍参加者の後に傍観者であったかの評価

- 傍参加者の後に話し手であったか

傍参加者の後に話し手であったかを図 6.9 に示す。グラフのエラーバーは標準偏差を示す。具体的には、傍参加者の後に受け手にならないまま話し手になった回数を評価する。Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。次に一元配置分散分析を行ったところ有意差が示されなかった ($p = 0.30$)。

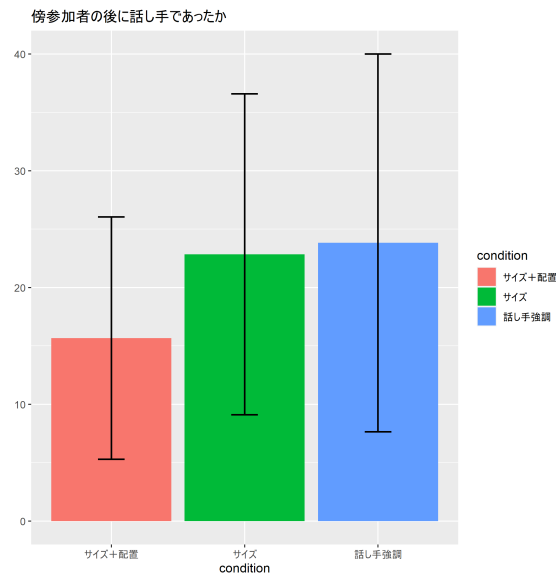


図 6.9: 傍参加者の後に話し手であったかの評価

- 傍観者の後に話し手であったか

傍観者の後に話し手であったかの評価を図 6.10 に示す。具体的には、傍観者の後に話し手になった回数を評価する。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意差が示されなかった ($p = 0.48$)。

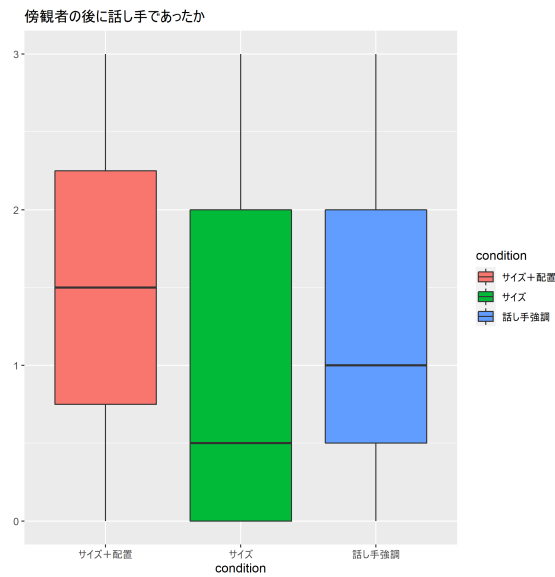


図 6.10: 傍観者の後に話し手であったかの評価

6.1.6 発話欲求度の平均

各条件における参加者の発話欲求度の平均を図 6.11 に示す。グラフのエラーバーは標準偏差を示す。発話欲求度の平均は、各参加者における 7 分間の会議において、計測した発話欲求度の和を計測した総フレーム数で割った数値である。各参加者の発話欲求度の平均に対して Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。次に一元配置分散分析を行ったところ有意差が示されなかった ($p = 0.96$)。各条件における平均値は参与役割提示 (サイズ+配置) 条件が 90.41, 参与役割提示 (サイズ) 条件が 90.70, 話し手強調表示条件が 90.10 であった。発話欲求度は 50 を閾値として傍参加者、傍観者を推定するため、この結果は参加者が傍観者になる状況がほとんど発生しなかったことを示唆している。

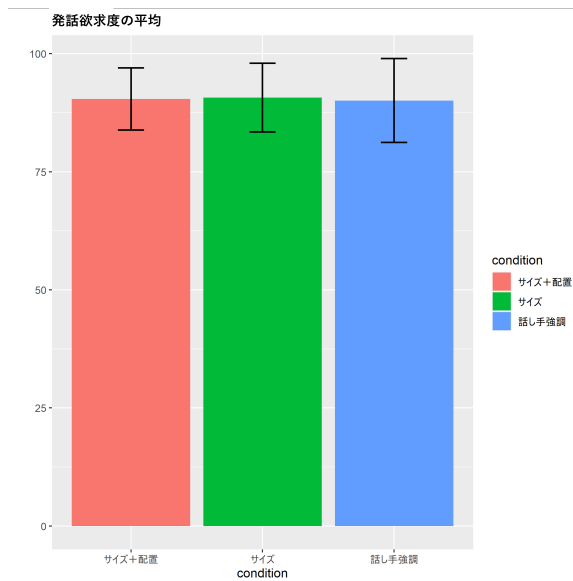


図 6.11: 各条件における参加者の発話欲求度の平均。

6.1.7 発話予備動作

発話予備動作を評価するために、話し手への視線配布した時間、頷き回数、挙手した回数および時間、開口した回数および時間を用いた。

- 話し手への視線配布した時間

各条件における参加者の話し手へ視線配布した時間の平均を図 6.12 に示す。グラフのエラーバーは標準偏差を示す。話し手への視線配布した時間は、各参加者における7分間の会議において、他の人が話し手である時にその人に視線配布を向けていた時間である。話し手への視線配布した時間に対して Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。次に一元配置分散分析を行ったところ有意差が示された ($p = 0.0074 < 0.010$)。Bonferroni-Holm の方法により有意水準を補正し、多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.0269 < 0.050$)。このことから、参与役割提示 (サイズ+配置) 条件は、話し手強調条件に比べて話し手へ視線配布を行っていることが示された。

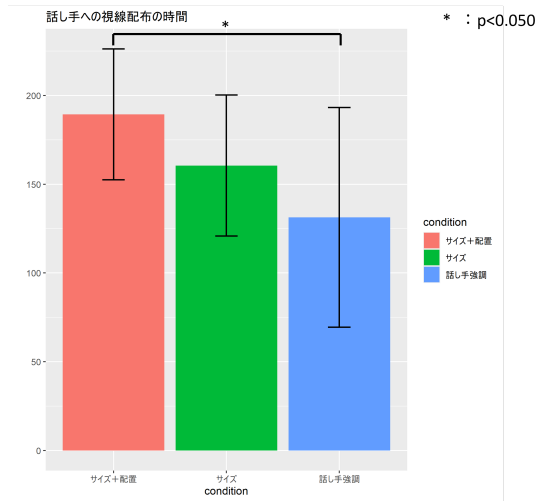


図 6.12: 各条件における参加者の話し手へ視線配布した時間.

- 頷き回数

各条件における参加者の頷きを行った回数を図 6.13 に示す. Shapiro-Wilk 検定を行った結果, 非正規性が示された. ART を行い, 一元配置分散分析を行ったところ, 有意差は示されなかった ($p = 0.88$).

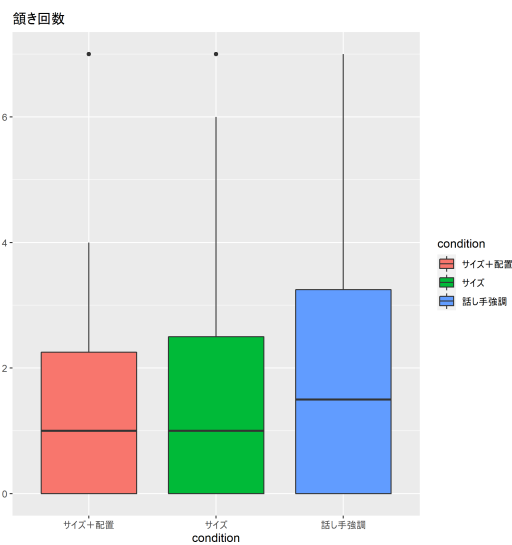


図 6.13: 各条件における参加者の頷きを行った回数.

- 挙手した回数および時間

各条件における参加者の挙手を行った回数を図 6.14 左に示す。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意差は示されなかった ($p = 0.42$)。また、各条件における参加者の挙手を行った時間を図 6.14 右に示す。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意差は示されなかった ($p = 0.28$)。

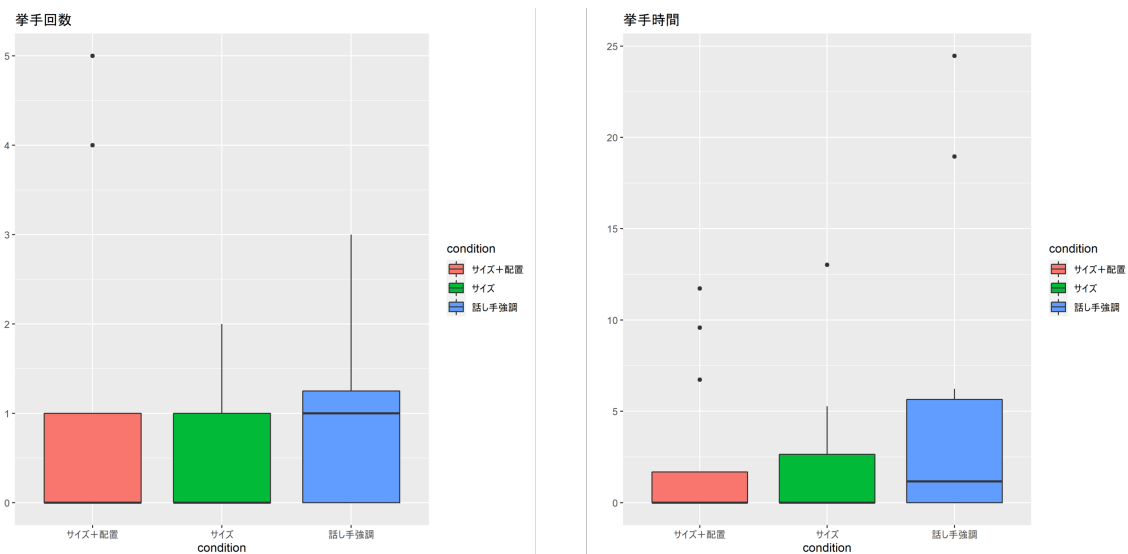


図 6.14: 各条件における参加者の挙手した回数（左）および時間（右）。

- 開口した回数および時間

各条件における参加者の開口を行った回数を図 6.15 左に示す。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意差は示されなかった ($p = 0.72$)。また、各条件における参加者の開口を行った時間を図 6.15 右に示す。Shapiro-Wilk 検定を行った結果、非正規性が示された。ART を行い、一元配置分散分析を行ったところ、有意差は示されなかった ($p = 0.75$)。

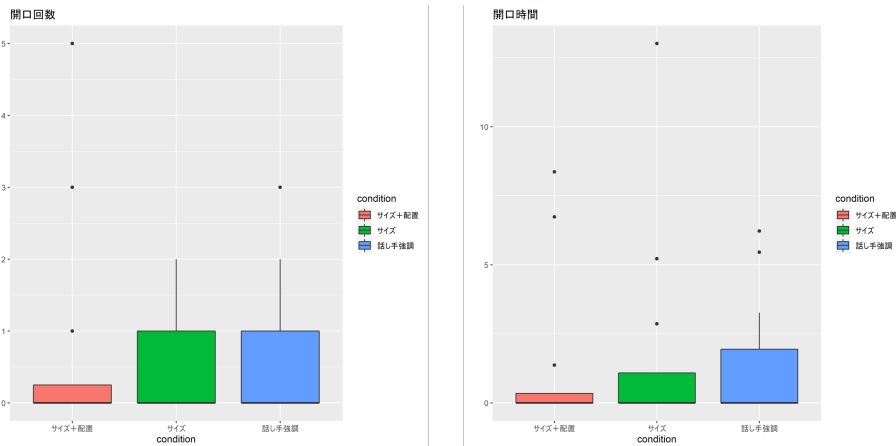


図 6.15: 各条件における参加者の開口した回数（左）および時間（右）。

6.2 アンケートによる印象評価

アンケートの各項目から印象評価を行う。

6.2.1 参与役割の把握

各条件間に行ったアンケート（表 5.2）のうち、参与役割の把握 (b) への印象に対する項目の評価を述べる。まず、それぞれの項目について Shapiro-Wilk 検定を行った結果、全ての項目において非正規性が示された。下記にそれぞれの項目について、アンケート項目が表す参与役割の把握に対する意図、および ART を用いて行った一元配置分散分析の結果について述べる。

- 今自分が会話に参加していること/していないことを意識することができた（図 6.16）

上記の項目について、自分が話者交替に関わる話し手、受け手、または傍参加者であるか、もしくは会話への参与が少ない傍観者・盗み聞き者であるかを把握できたかについての印象評価を行うための項目である。5段階リッカート尺度を用いて得られた回答に対して ART を行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.015 < 0.050$)。多重比較を行ったところ、参与役割提示（サイズ+配置）条件-参与役割提示（サイズ）条件間において有意差が示された ($p = 0.042 < 0.050$)。また、参与役割提示（サイズ+配置）条件-話し手強調提示条件間において有意差が示された ($p = 0.021 < 0.050$)。このことから、参与役割提示（サイズ+配置）条件は、参与役割提示（サイズ）条件および話し手強調提示条件に対して、自分が話者交替に関わる参与役割かどうかを把握しやすいと参加者が感じていることが示された。

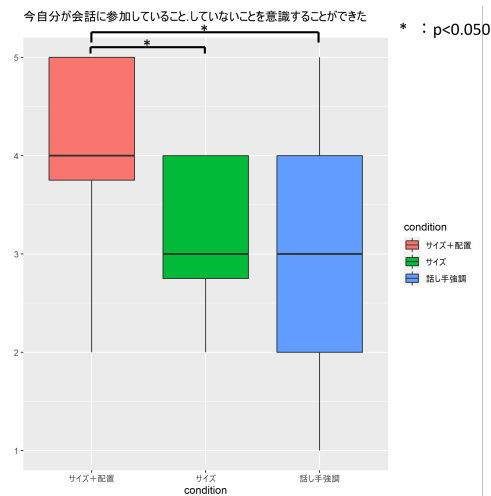


図 6.16: 各条件における b-1 「今自分が会話に参加していること/していないことを意識することができた」に対する回答.

- 自分の話を聞いている参加者がわかった (図 6.17)

上記の項目について、自分が話し手である時、他の参加者のうち誰が受け手および傍参加者であるか把握できたかについての印象評価を行うための項目である。5段階リッカート尺度を用いて得られた回答に対して ART を行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.028 < 0.050$)。多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-参与役割提示 (サイズ) 条件間において有意差が示された ($p = 0.025 < 0.050$)。このことから、参与役割提示 (サイズ+配置) 条件は、参与役割提示 (サイズ) 条件に対して、自分が話し手の時に誰と話者交替を行うかを把握しやすいと参加者が感じていることが示された。

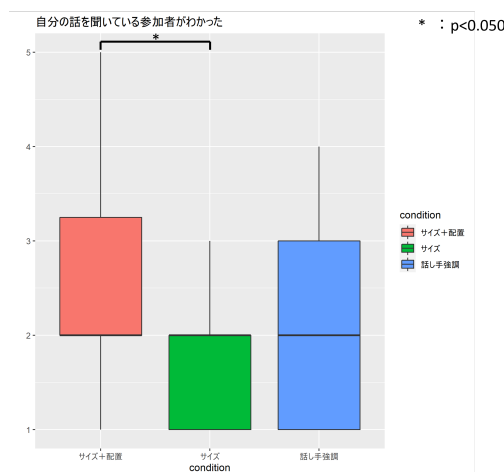


図 6.17: 各条件における b-2 「自分の話を聞いている参加者がわかった」に対する回答.

- 自分がちゃんと聞いていることが相手に伝わっているように感じた (図 6.18)

上記の項目について、自分が傍参加者である時、話し手に対して自分が傍参加者であることが提示されていることを把握できたかについての印象評価を行うための項目である。5段階リッカート尺度を用いて得られた回答に対してARTを行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.028 < 0.050$)。多重比較を行ったところ、参与役割提示(サイズ+配置)条件-話し手強調提示条件間において有意差が示された ($p = 0.022 < 0.050$)。このことから、参与役割提示(サイズ+配置)条件は、話し手強調提示条件に対して、自分が傍参加者であることが話し手に提示されていることを把握しやすいと参加者が感じていることが示された。

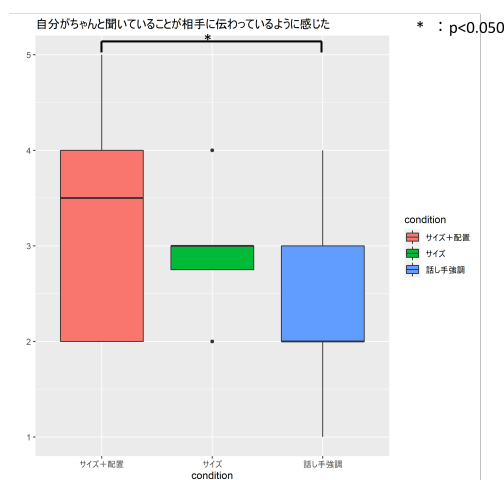


図 6.18: 各条件における b-3 「自分がちゃんと聞いていることが相手に伝わっているように感じた」に対する回答。

- 自分が話し終えるとき、次に話たさそうな人がわかった (図 6.19)

上記の項目について、自分が話し手である時、受け手となりうる参加者を把握できたかについての印象評価を行うための項目である。5段階リッカート尺度を用いて得られた回答に対してARTを行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.018 < 0.050$)。多重比較を行ったところ、参与役割提示(サイズ+配置)条件-話し手強調提示条件間において有意差が示された ($p = 0.014 < 0.050$)。このことから、参与役割提示(サイズ+配置)条件は、話し手強調提示条件に対して、自分が話し手である時に受け手となりうる参加者を把握しやすいと参加者が感じていることが示された。

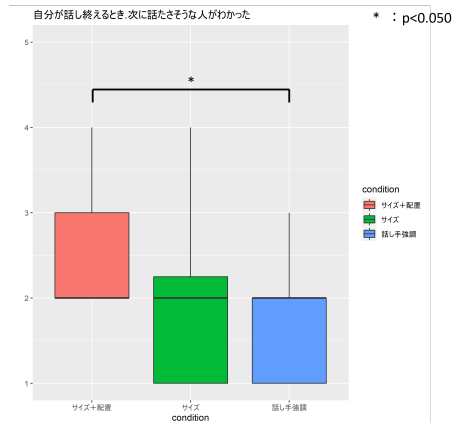


図 6.19: 各条件における b-4 「自分が話し終えるとき、次に話たさそうな人がわかった」に対する回答.

- 自分が話し終えるとき、話したくない人がわかった (図 6.20)

上記の項目について、自分が話し手である時、受け手とならない傍観者である参加者を把握できたかについての印象評価を行うための項目である。5段階リッカート尺度を用いて得られた回答に対して ART を行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.014 < 0.050$)。多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-参与役割提示 (サイズ) 条件間において有意差が示された ($p = 0.047 < 0.050$)。また、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.018 < 0.050$)。このことから、参与役割提示 (サイズ+配置) 条件は、参与役割提示 (サイズ) 条件および話し手強調提示条件に対して、自分が話し手である時に傍観者である参加者を把握しやすいと参加者が感じていることが示された。

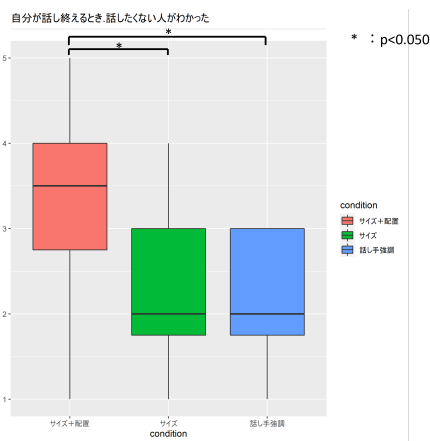


図 6.20: 各条件における b-5 「自分が話し終えるとき、話したくない人がわかった」に対する回答.

- 他の人が話し終わるとき、自分が話すべきかがわかった (図 6.21)

上記の項目について、話し手である他の参加者との話者交替時に、自分が受け手であることを把握できたかについての印象評価を行うための項目である。5段階リッカート尺度を用いて得られた回答に対してARTを行い、一元配置分散分析を行った結果、有意差が示された ($p = 0.045 < 0.050$)。多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-話し手強調提示条件間において有意差が示された ($p = 0.036 < 0.050$)。このことから、参与役割提示 (サイズ+配置) 条件は、話し手強調提示条件に対して、話者交替時に自分が受け手であることが把握しやすいと参加者が感じていることが示された。

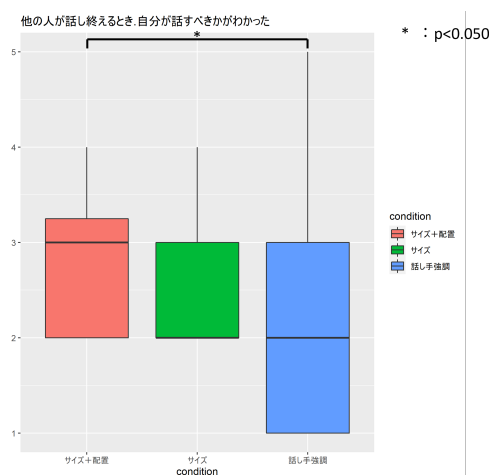


図 6.21: 各条件における b-6 「他の人が話し終わるとき、自分が話すべきかがわかった」に対する回答。

6.2.2 話者交替および会議への評価

各条件間に行ったアンケート (表 5.2) のうち、会議および話者交替への印象 (c) に対する項目の評価を述べる。まず、それぞれの項目について Shapiro-Wilk 検定を行った結果、全ての項目において非正規性が示された。下記にそれぞれの項目について、ART を用いて行った一元配置分散分析の結果について述べる。

- 議論に集中できた (図 6.22)

5段階リッカート尺度を用いて得られた回答に対してARTを行い、一元配置分散分析を行った結果、有意差が示されなかった ($p = 0.17$)。

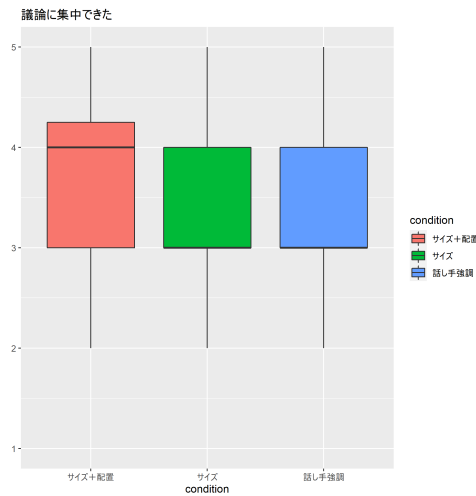


図 6.22: 各条件における c-1 「議論に集中できた」に対する回答.

- 議論にストレスを感じた (図 6.23)

5段階リッカート尺度を用いて得られた回答に対して ART を行い、一元配置分散分析を行った結果、有意傾向が示された ($p = 0.073 < 0.10$). 多重比較を行ったところ、参与役割提示 (サイズ+配置) 条件-参与役割提示 (サイズ) 提示条件間において有意傾向が示された ($p = 0.060 < 0.10$). このことから参与役割提示 (サイズ+配置) 条件は、参与役割提示 (サイズ) 提示条件に対して議論へのストレスが軽減される可能性が示唆された.

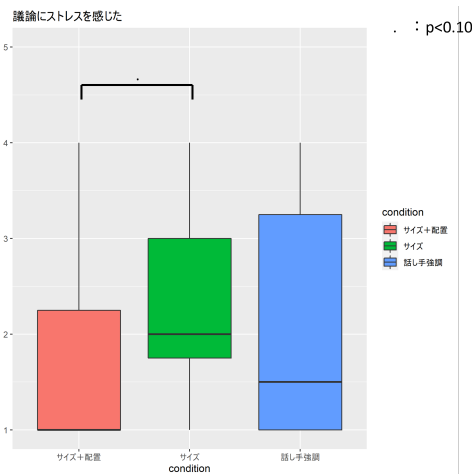


図 6.23: 各条件における c-2 「議論にストレスを感じた」に対する回答.

- 議論に疲労を感じた (図 6.24)

5段階リッカート尺度を用いて得られた回答に対してARTを行い、一元配置分散分析を行った結果、有意差が示されなかった ($p = 0.35$).

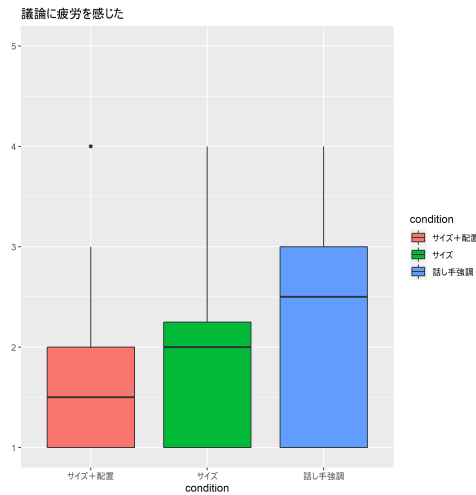


図 6.24: 各条件における c-3 「議論に疲労を感じた」に対する回答.

- 話者交替にストレスを感じた (図 6.25)

5段階リッカート尺度を用いて得られた回答に対してARTを行い、一元配置分散分析を行った結果、有意差が示されなかった ($p = 0.55$).

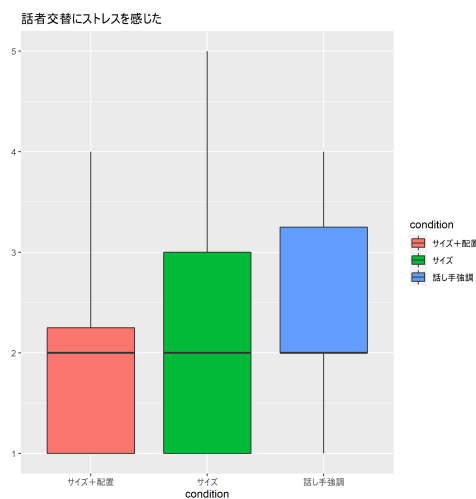


図 6.25: 各条件における c-4 「話者交替にストレスを感じた」に対する回答.

6.2.3 AUTによる創造性の評価

各条件間に行ったアンケート（表 5.2）のうち，AUTによる創造性に対する評価 (a) を述べる。

- アイデアの数はたくさん考えられたか（図 6.26）

これは流暢性を評価する項目である。Shapiro-Wilk 検定を行った結果，非正規性が示された。ART を行い，一元配置分散分析を行った結果，有意差は示されなかった ($p = 0.91$)。

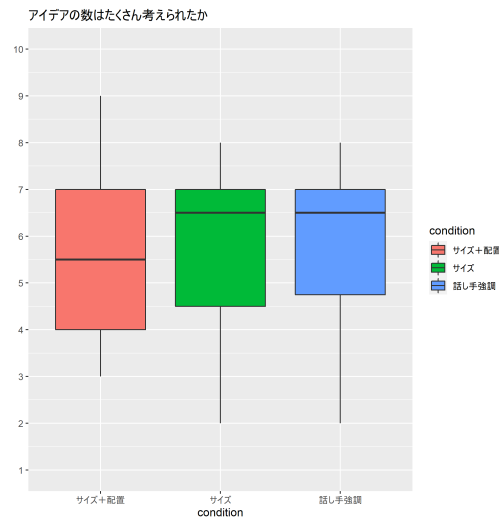


図 6.26: 各条件における a-1 「アイデアの数はたくさん考えられたか」に対する回答。

- 独創的なアイデアを考えることができたか（図 6.27）

これは独創性を評価する項目である。Shapiro-Wilk 検定を行った結果，非正規性が示された。ART を行い，一元配置分散分析を行った結果，有意差は示されなかった ($p = 0.99$)。

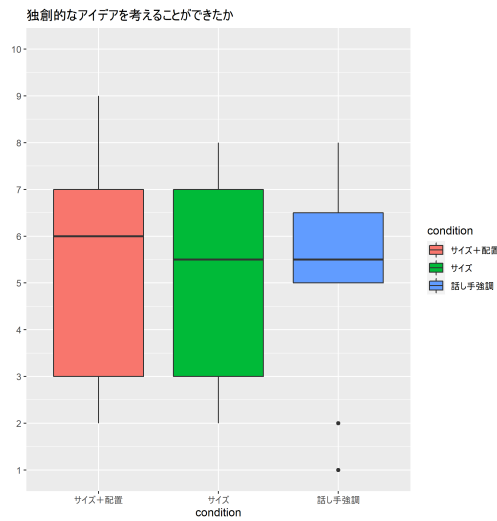


図 6.27: 各条件における a-2 「独創的なアイデアを考えることができたか」に対する回答.

- 幅広い用途のアイデアを考えることができたか (図 6.28)

これは柔軟性を評価する項目である. グラフのエラーバーは標準偏差を示す. Shapiro-Wilk 検定を行った結果, 正規性が示された. そのため, Bartlett 検定を行った結果, 等分散性が示された. 一元配置分散分析を行った結果, 有意差は示されなかった ($p = 0.98$).

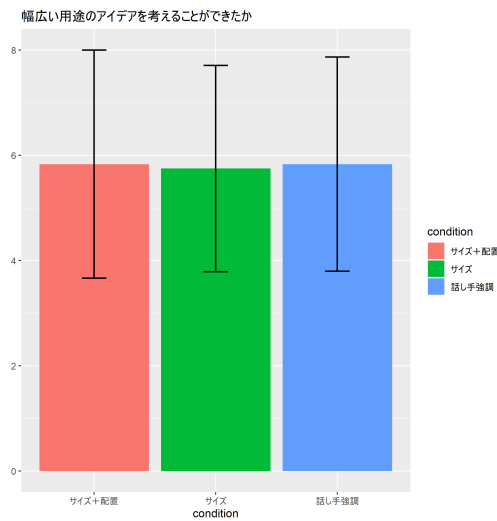


図 6.28: 各条件における a-3 「幅広い用途のアイデアを考えることができたか」に対する回答.

- 詳細なアイデアを考えることができたか (図 6.29)

これは綿密性を評価する項目である。グラフのエラーバーは標準偏差を示す。Shapiro-Wilk 検定を行った結果、正規性が示された。そのため、Bartlett 検定を行った結果、等分散性が示された。一元配置分散分析を行った結果、有意差は示されなかった ($p = 0.23$)。

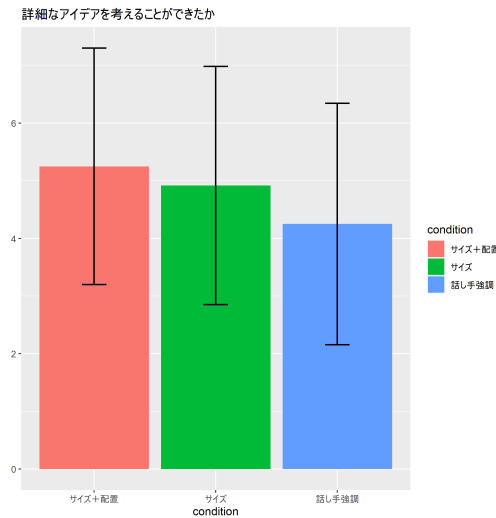


図 6.29: 各条件における a-4 「詳細なアイデアを考えることができたか」に対する回答。

6.3 考察

本節では実験の結果を踏まえて、各仮説に対する考察を行う。

6.3.1 仮説 1：参与役割提示条件では、話し手強調提示条件と比較して、参与役割の把握が促進される。

参与役割提示条件において、推定した参与役割を提示することから参与役割の把握が促進されると考えられる。実験の結果から、参与役割の遷移について、全ての項目で条件間に有意差は示されなかった。この原因について、参与役割の推定精度が不足していたことから、設計意図と異なり、傍観者から傍参加者への参与役割の遷移が発生しにくかったからであると考えられる。本実験では参与役割の把握を評価する補足として、発話欲求度の平均、および発話予備動作を用いて参加者の行動から評価した。発話予備動作として、話し手への視線配布の時間、挙手した回数および時間、頷き回数、および開口した回数および時間を評価した。発話欲求度の平均について、条件間に有意差は示されず、全ての条件について平均値は 90 以上であった。発話欲求度は 50 を超えることにより傍観者から傍参加者へ遷移する。そのため、平均値が 90 以上であることから会議においてほとんど傍観者と推定されることはなかったと

考えられる。また、挙手した回数および時間、頷き回数、および開口した回数および時間について条件間に有意差は示されなかった。これは、会議において発話欲求度が50を下回ることが少なく、意識的に発話予備動作を行い参与役割を傍観者から傍参加者へ遷移させる必要がほとんどなかったためであると考えられる。以上の理由から、傍観者から傍参加者または傍参加者から傍観者への参与役割の遷移が発生しにくく、提示された参与役割を意識した話者交替が困難であったと考えられる。一方で話し手への視線配布の時間について、参与役割提示（サイズ+配置）条件は、話し手強調条件に比べて話し手へ視線配布を有意に多く行っていることが示された。これは参与役割提示（サイズ+配置）条件において、話者交替が行われる空間が黄色の円で強調されており、視線配布を送る話し手が把握しやすかったためであると考えられる。

次に、参与役割の把握が促進されたかを定性的に評価するため、アンケートにより参与役割の把握に対する印象を評価した。アンケートの結果より、参与役割提示（サイズ+配置）条件は、参与役割提示（サイズ）条件および話し手強調提示条件に対して、「自分が話者交替に関わる参与役割かどうか」および「自分が話し手である時に傍観者である参加者」を把握しやすいと参加者が感じていることが示された。これは参与役割提示（サイズ+配置）条件において、話者交替が行われる空間が黄色の円で強調されているため、ビデオウィンドウが円の内部に表示されているかによって参与役割を把握することが可能であるからだと考えられる。また、参与役割提示（サイズ+配置）条件は、参与役割提示（サイズ）条件に対して、「自分が話し手の時に誰と話者交替を行うか」を把握しやすいと参加者が感じていることが示された。これは参与役割提示（サイズ+配置）条件が話者交替に参加する参与役割が配置によって提示されている一方で、参与役割提示（サイズ）条件はウィンドウサイズのみで提示されるため、話者交替に参加する参与役割の把握が参与役割提示（サイズ+配置）条件に比べて困難であるためだと考えられる。最後に、参与役割提示（サイズ+配置）条件は、話し手強調提示条件に対して、「自分が傍参加者であることが話し手に提示されていること」、「自分が話し手である時に受け手となりうる参加者」、および「話者交替時に自分が受け手であること」を把握しやすいと参加者が感じていることが示された。これは参与役割提示（サイズ+配置）条件において、4つの参与役割を提示するのに対し、話し手強調提示条件は話し手および話し手以外の2つの役割のみ提示するためであると考えられる。また、参与役割提示（サイズ）条件について、参与役割提示（サイズ+配置）条件と同様に4つの参与役割を提示するが、ウィンドウサイズのみを用いて提示するため、会議中に参与役割の遷移を明確に把握することが困難であると考えられる。そのため、話し手強調条件に対して有意差が示されなかったと考えられる。

以上の結果をまとめると、定量的な評価（参与役割の遷移の評価）においては、参与役割提示条件で話し手強調提示条件よりも参与役割の把握が促進されるという結果は得られなかった。一方、定性的な評価（アンケートによる印象評価）においては、参与役割提示（サイズ+配置）条件で話し手強調提示条件よりも参与役割の把握に対する印象が向上することが示された。このことから仮説は一部支持されたと考えられる。つまり、参与役割提示条件について、提示された参与役割を把握することは促進される一方、提示された参与役割を意識した

話者交替が困難であった。この課題に対し、参与役割の推定精度を上げることで提示された参与役割を意識した話者交替が可能であると考えられる。

6.3.2 仮説2：参与役割提示条件では話し手強調提示条件と比較して、円滑な話者交替が促進される。

参与役割提示条件では話し手強調提示条件と比較して、参与役割の把握が促進されることにより、円滑な話者交替が促進されると考えられる。実験の結果から発話数について、一元配置分散分析の結果有意差が示されたが、条件間に有意差は示されなかった。一方で、話者交替数について、話し手強調提示条件は参与役割提示条件に対して、話者交替数が有意に多いことが示された。また話者交替の失敗が軽減したかを評価するため、発話衝突の回数および確率、沈黙時間を評価した。発話衝突の回数および確率において参与役割提示（サイズ+配置）条件は話し手強調提示条件に比べ、回数は有意に小さく、確率も有意に低かった。また自分が発話している際に他の人が発話を開始した際の発話衝突の回数について、参与役割提示（サイズ）条件は話し手強調提示条件に比べ、有意に回数が小さかった。このことから話し手強調提示条件において話者交替数が有意に多く、同時に発話衝突も有意に多いことが示された。これは発話衝突が多いことで発話数が多くなったと考えられる。以上より、話し手強調提示条件において円滑な話者交替が行われなかったと考えられる。また、参与役割提示（サイズ+配置）条件は発話衝突を減少させることが示された。この理由について、参与役割提示（サイズ+配置）条件において、次に話すべき受け手を把握することが可能であるためであると考えられる。沈黙について、条件間に有意差は示されなかった。これは分析において、アイデアを考えるための沈黙と話者交替の失敗による沈黙を区別していないためである可能性がある。

発話衝突の回数および確率に有意差が示された、一方話者交替へのストレスについて印象評価に有意差がなかった。その理由として、話し手強調提示条件について発話衝突が発生しストレスを感じる一方、参与役割提示条件において参与役割によって話したくても話しにくい状況がストレスを感じたのではないかと考えられる。

参与役割提示（サイズ+配置）条件は話し手強調提示条件に比べ、話者交替数は有意に少なく、発話衝突の回数および確率も有意に少なかった。このことから、話し手強調条件は参与役割の把握が難しく、話者交替が多い分、発話衝突が多くなると考えられる。一方参与役割提示（サイズ+配置）条件は会議において参与役割を把握して話者交替を行うため、発話衝突が減少したと考えられる。以上のことから仮説は部分的に支持されたと考えられる。

6.3.3 仮説3：参与役割提示条件は話し手強調提示条件と比較して、会議の質が向上する。

円滑な話者交替が促進することにより、会議への負荷が減少し、会議の質が向上すると考えられる。実験の結果から、議論に集中できたかについて参加者に回答させた評価において条件に有意差はなかった。議論へストレスを感じたかについて参与役割提示（サイズ+配置）条件は、参与役割提示（サイズ）提示条件に対して議論へのストレスが軽減される可能性が示唆された。これは参与役割提示（サイズ）提示条件について4つの参与役割をウィンドウサイズのみを用いて表現するため、参与役割提示（サイズ+配置）条件に比べて参与役割の把握が困難である。そのため、参与役割を把握しながら行う議論へストレスを感じたのではないかと考える。議論に疲労を感じたかについて参加者に回答させた評価において条件に有意差はなかった。議論に集中できたか、および議論に疲労を感じたかについて有意差が現れなかった原因として、会議時間の短さが挙げられる。本研究では時間経過による会議への質を落とさないため、1つの議論を7分に設定した。そのため、議論への集中および疲労に差が示される前に会議が終了してしまった可能性が考えられる。

また、アンケートからAUTによる創造性への影響を評価した。創造性を、流暢性、独創性、柔軟性、および綿密性から評価した。しかし、全ての項目において、有意差は示されなかった。一方で、題材について評価（図 6.30）を行ったところ、綿密性について有意差が示された（ $p = 0.031 < 0.050$ ）。多重比較を行ったところ、靴下-鍋間（ $p = 0.039 < 0.050$ ）に有意差が示された。これは題材の考えやすさが創造性に影響することを示している。そのため、条件間の影響を調査するためには、題材の考えやすさを考慮した実験設計が必要であると考えられる。

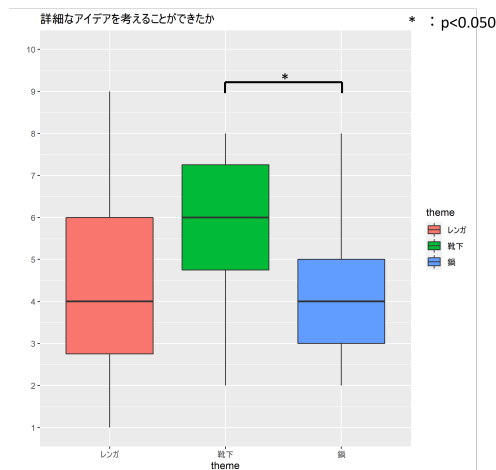


図 6.30: 各題材に対する綿密性の評価.

第7章 今後の課題

本章では実験で得た考察をもとに今後の課題を述べる。

7.1 参与役割の推定手法について

本実験において、参与役割の遷移について有意差が示されなかった。この原因について、参与役割を推定する精度が不足していたことが挙げられる。参与役割を推定する精度をあげるために、今後の課題として2つの改善点を述べる。

まず、発話欲求度の算出に使われる発話予備動作の選択である。本研究において、発話欲求度の算出のために、話し手への視線配布、頷き、挙手、開口、および直前の発話を用いた。しかし、本研究においてこれらの発話予備動作はほとんど見られなかった。実験後に行ったシステムの印象について記述式で回答させたアンケートにおいても、挙手をほとんど使わなかったという意見が多くみられた。また「視線の部分が意識的に顔を向けなければならないので、あまり利用できなかった気がした。」といった意見も見られた。以上のことから、本研究において選択した発話予備動作はビデオ会議における使用の負荷が大きい可能性がある。よって今後の改善点として、ビデオ会議システムにおいて自然に発話欲求が表出する発話予備動作を調査することで参与役割の推定する精度が向上する可能性が考えられる。

次に、発話欲求度の増減量の調整である。本研究において、過去の研究を参考に著者が独自に発話欲求度の増減量を調整した。しかし、本実験において、発話欲求度はすべての条件において90を超えていた。これにより、参加者は傍観者になることがほとんどなかった。以上のことから、参加者の発話欲求を十分に反映できなかった可能性が考えられる。よって今後の改善点として、発話予備動作に対する十分に発話欲求を反映することが可能な増減量を調査することで参与役割の推定する精度が向上する可能性が考えられる。

7.2 参与役割の提示手法について

本実験における参与役割提示（サイズ+配置）条件について、ウィンドウサイズおよび配置を用いて参与役割を提示した。また、参与役割提示（サイズ）条件について、ウィンドウサイズのみを用いて参与役割を提示した。アンケートによる参与役割の把握の印象評価において、参与役割提示（サイズ+配置）条件について話し手強調提示条件と比較して有意差が示された一方、参与役割提示（サイズ）条件については話し手強調提示条件と比較して有意差が示されなかった評価があった。また、実験後に行ったシステムの印象について記述式で

回答させたアンケートにおいて、「参与役割提示（サイズ+配置）条件はわかりやすかったが参与役割提示（サイズ）条件は誰が話したいかわからなかった」といった意見や「ウィンドウのサイズはあまり見る余裕がなかった」といった意見が見られた。このことから配置が参与役割を把握を促進し、ウィンドウサイズによる参与役割の提示が機能していなかった可能性が考えられる。これに対し、今後の課題としてウィンドウサイズによる参与役割の提示における大きさの調整が挙げられる。調整には、どの参与役割かが即座に明確に把握できる大きさに設計する必要があると考えられる。また、実験後に行ったシステムの印象について記述式で回答させたアンケートにおいて、「ユーザの顔を映す場所が小さかった」といった意見が見られた。このことから、明確に参与役割を把握することが可能なだけでなく、ビデオ会議システムとして表情が把握できることも考慮する必要があると考えられる。

また本研究で実装したシステムにおいて、参与役割が遷移しビデオウィンドウが変化する際、一度消え、その後参与役割が遷移した後のビデオウィンドウが表示される。そのため、実験後に行ったシステムの印象について記述式で回答させたアンケートにおいて、ビデオウィンドウが切り替わる際の点滅が気になるといった旨の意見が見られた。これに対し、今後の改善点として、ビデオウィンドウが一度消えることなく、連続的に変化していくことにより、点滅による会議への障害が軽減されることが考えられる。

7.3 実験設計について

会議を促進したかを評価するためにAUTによる創造性の評価を行った。AUTにおいて、創造性を流暢性、独創性、柔軟性、および綿密性の観点から評価する。ここで綿密性について、お題間に有意差が示された。つまり、お題による影響が綿密性に示された。しかし、本実験において、条件間による創造性を評価する必要があった。そのため、今後の課題としてそれぞれのお題におけるAUTの会議のしやすさを考慮して実験設計する必要があると考えられる。

また本実験の参加者は、12人2グループで評価を行った。順序効果を避けるため、グループごとに行う条件の順序を変えたが、実験の評価に条件の順序による影響が発生した可能性がある。そのため、今後の改善点として、順序による影響がなくなるような実験設計が必要である。

第8章 おわりに

本研究では、多人数ビデオ会議において、話者交替を促進することを目的に参与役割を推定し、提示する手法を提案した。著者は先行研究において、視線配布から3つの参与役割を推定し、ウィンドウサイズの大きさとして提示する手法を提案した。しかし、参与役割の推定精度が不十分であることや参与役割の能動的な制御が困難であること等の課題が示唆された。これに対し、本研究において会議参加者を話し手、受け手、傍参加者、傍観者の4つの参与役割に推定する。参与役割の推定には、各参加者が会議中に行う非言語情報を用いた。推定した参与役割はビデオウィンドウのウィンドウサイズおよび配置を用いて参与役割を提示した。本研究では提案手法を実装したビデオ会議システムが実際に話者交替を促進するかを調査することを目的に実験を行った。実験において、6人1グループで創造会議を行った。実験条件として、参与役割提示（サイズ+配置）条件、参与役割提示（サイズ）条件、および話し手強調提示条件を比較した。実験の結果、参与役割の把握を促進する可能性があることが明らかになった。また、発話衝突を減少させることが示された。一方で参与役割の推定する精度が低いことやウィンドウサイズによる提示が機能していない可能性があることなどの課題が明らかになった。今後の展望として、提案手法に適した発話予備動作の選択や参与役割が明確に把握することが可能な提示手法について検討することが挙げられる。

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、川口一画先生には多大な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、実質指導教員である川口一画先生には、研究の進め方、論文執筆、発表の方法といった研究に関することのみならず、日常生活における研究者の心構えなど多岐に渡るご指導を頂きました。先生の熱心なご指導のおかげで、これまでの研究成果の発表および本論文の執筆ができました。ここに深く感謝の意を表します。

インタラクティブプログラミング研究室の後輩、同輩、先輩方には研究生活においてお世話になりました。特に COMMUNICATION 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多くのご支援に加えて、研究室においても研究の進め方や研究生活について多くのご助言を頂きました。研究室の皆様のおかげで、3年間充実した研究生活を送ることができました。深く感謝いたします。

最後に、学生生活においてお世話になった皆様、そして、私の学生生活を支えて頂いた家族に深く感謝を申し上げます。

参考文献

- [20008] 会議の進め方. 日経文庫. 日本経済新聞出版社, 2008.
- [ACH96] Russell L Andersson, Tsuhan Chen, and Barin G Haskell. Video conference system and method of providing parallax correction and a sense of presence, 1996.
- [AMH⁺] Amin G. Alhashim, Megan Marshall, Tess Hartog, Rafal Jonczyk, Danielle Dickson, Janet van Hell, GÅ¼l E. Okudan-Kremer, and Zahed Siddique. Wip: Assessing creativity of alternative uses task responses: A detailed procedure.
- [Bir10] Ray L Birdwhistell. *Kinesics and context: Essays on body motion communication*. University of Pennsylvania press, 2010.
- [Cor] NTT Communications Corporation. Skyway. <https://webrtc.ecl.ntt.com/>.
- [Cre22] CreativeHuddle. The alternative uses test, 2022. <https://www.creativehuddle.co.uk/post/the-alternative-uses-test>.
- [Dav73] James H Davis. Group decision and social interaction: A theory of social decision schemes. 1973.
- [Dis] Discord. Discord. <https://discord.com/>.
- [Eva22] Bob Evans. The zoom revolution: 10 eye-popping stats from tech’s new superstar, Nov 2022.
- [GM03] David M. Grayson and Andrew F. Monk. Are you looking at me? eye contact and desktop video conferencing. 2003.
- [Gof76] Erving Goffman. Replies and responses. *Language in Society*, Vol. 5, No. 3, pp. 257–313, 1976.
- [Goo] *Conversational Organization: Interaction Between Speakers and Hearers*.
- [GP] Inc. Gather Presence. Gather. <https://www.gather.town/>.
- [GS07] Christopher A Gorse and Anne Marie Sanderson. Exploring group work dynamics. In *Proceedings of the 23rd Annual ARCOM Conference*, pp. 3–5, 2007.

- [HAH22] Erzhen Hu, Md Aashikur Rahman Azim, and Seongkook Heo. Fluidmeet: Enabling frictionless transitions between in-group, between-group, and private conversations during virtual breakout meetings. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [Hal92] E.T. Hall. *The Hidden Dimension*. Anchor Books, 1992.
- [HBT90] James J Higgins, R Clifford Blair, and Suleiman Tashtoush. The aligned rank transform procedure. 1990.
- [IKST21] Rikuto Iitsuka, Ikkaku Kawaguchi, Buntarou Shizuki, and Shin Takahashi. Multi-party video conferencing system with gaze cues representation for turn-taking. In Davinia Hernández-Leo, Reiko Hishiyama, Gustavo Zurita, Benjamin Weyers, Alexander Nolte, and Hiroaki Ogata, editors, *Collaboration Technologies and Social Computing*, pp. 101–108, Cham, 2021. Springer International Publishing.
- [Ken67] Adam Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, Vol. 26, pp. 22 – 63, 1967.
- [Ken73] Adam Kendon. The role of visible behavior in the organization of social interaction. *Social communication and movement*, pp. 29–74, 1973.
- [KKV⁺11] Jan GM Kooloos, Tim Klaassen, Mayke Vereijken, Sascha Van Kuppeveld, Sanneke Bolhuis, and Marc Vorstenbosch. Collaborative group work: Effects of group size and assignment structure on learning gain, student satisfaction and perceived participation. *Medical Teacher*, Vol. 33, No. 12, pp. 983–988, 2011.
- [KKZ⁺17] Kazuki Kobayashi, Takashi Komuro, Bo Zhang, Keiichiro Kagawa, and Shoji Kawahito. A gaze-preserving group video conference system using screen-embedded cameras. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, VRST '17, New York, NY, USA, 2017. Association for Computing Machinery.
- [KPA22] Katherine A Karl, Joy V Peluchette, and Navid Aghakhani. Virtual work meetings during the covid-19 pandemic: The good, bad, and ugly. *Small Group Research*, Vol. 53, No. 3, pp. 343–365, 2022.
- [LTN⁺19] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, Wan-Teh Chang, Wei Hua, Manfred Georg, and Matthias Grundmann. Mediapipe: A framework for building perception pipelines, 2019.

- [Mic] Microsoft. Microsoft teams. <https://www.microsoft.com/ja-jp/microsoft-365/microsoft-teams/free>.
- [MOAY05] Naoki Mukawa, Tsugumi Oka, Kumiko Arai, and Masahide Yuasa. What is connected by mutual gaze? user's behavior in video-mediated communication. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, p. 1677–1680, New York, NY, USA, 2005. Association for Computing Machinery.
- [MP59] Arnold Meadow and Sidney J Parnes. Evaluation of training in creative problem solving. *Journal of Applied Psychology*, Vol. 43, No. 3, p. 189, 1959.
- [NYS⁺14] Naoto Nakazato, Shigeo Yoshida, Sho Sakurai, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Smart face: Enhancing creativity during video conferences using real-time facial deformation. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work Social Computing*, CSCW '14, p. 75–83, New York, NY, USA, 2014. Association for Computing Machinery.
- [OMIM94] Ken-Ichi Okada, Fumihiko Maeda, Yusuke Ichikawaa, and Yutaka Matsushita. Multiparty videoconferencing at virtual social distance: Majic design. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, CSCW '94, pp. 385—393, New York, NY, USA, 1994. Association for Computing Machinery.
- [Spa] SpatialChat. Spatialchat. <https://www.spatial.chat/>.
- [SSJ78] Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*, pp. 7–55. Elsevier, 1978.
- [Var86] Marjorie Fink Vargas. *Louder than words : an introduction to nonverbal communication*. Iowa State University Press, 1986.
- [Ver99] Roel Vertegaal. The gaze groupware system: Mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pp. 294—301, New York, NY, USA, 1999. Association for Computing Machinery.
- [VSvdVN01] Roel Vertegaal, Robert Slagter, Gerrit van der Veer, and Anton Nijholt. Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pp. 301—308, New York, NY, USA, 2001. Association for Computing Machinery.
- [VVdVV00] Roel Vertegaal, Gerrit Van der Veer, and Harro Vons. Effects of gaze on multiparty mediated communication. In *Graphics interface*, pp. 95–102, 2000.

- [VWSC03] Roel Vertegaal, Ivo Weevers, Changuk Sohn, and Chris Cheung. Gaze-2: Conveying eye contact in group video conferencing using eye-controlled camera direction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, p. 521–528, New York, NY, USA, 2003. Association for Computing Machinery.
- [WFGH11] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, p. 143–146, New York, NY, USA, 2011. Association for Computing Machinery.
- [win20] Microsoft teams hits 75 million daily active users, up from 44 million in march, Apr 2020.
- [WLB00] James R Wilcox, Laura Long, and Michel L Bayard. *Videoconferencing & interactive multimedia: the whole picture*. Routledge, 2000.
- [WLV⁺21] Julie Williamson, Jie Li, Vinoba Vinayagamoorthy, David A Shamma, and Pablo Cesar. Proxemics and social interactions in an instrumented virtual reality workshop. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.
- [ZVC] Inc. Zoom Video Communications. Zoom. <https://zoom.us/>,.
- [井上 99] 井上智雄, 岡田謙一, 松下温. テレビ会議における映像表現の利用とその影響. 情報処理学会論文誌, 1999.
- [玉木 12] 玉木秀和, 東野豪, 小林稔, 井原雅行, 岡田謙一. 遠隔会議における発話衝突低減手法. 情報処理学会論文誌, 2012.
- [玉木 13] 玉木秀和, 東野豪, 小林稔, 井原雅行. 音声遅延が遠隔会議中の発話衝突と精神的ストレスに与える影響. 電子情報通信学会論文誌 D, Vol. 96, No. 1, pp. 35–45, 2013.
- [原岡 90] 原岡一馬. 人間とコミュニケーション. ナカニシヤ出版, 1990.
- [高橋 87] 高橋誠. 会議の進め方. 日本経済新聞出版社, 1987.
- [佐藤 11] 佐藤良, 坂本孝丈, 兵藤幸与, 石川智之, 市川淳, 後藤知彦, 竹内勇剛. 円滑な会話を実現する多人数ビデオチャット環境の構築. 2011 年度日本認知科学会第 28 回大会予稿集, Vol. 2019, pp. 205–212, 2011.

- [佐藤 13] 佐藤良, 竹内勇剛. 多人数対話におけるロボットの視線行動に基づく発話権と対話場のデザイン. HAI シンポジウム 2013, pp. 219–228, 2013.
- [山岡 17] 山岡明奈, 湯川進太郎. マインドワンダリングおよびアウェアネスと創造性の関連. 社会心理学研究, Vol. 32, No. 3, pp. 151–162, 2017.
- [山田 22] 山田楓也, 石田繁巳, 白石陽. Web 会議におけるミクロな顔特徴に着目した発話予測手法の検討. Technical Report 19, 公立はこだて未来大学大学院システム情報科学研究科, 公立はこだて未来大学システム情報科学部, 公立はこだて未来大学システム情報科学部, mar 2022.
- [小磯 00] 小磯花絵, 伝康晴. 円滑な話者交替はいかにして成立するか. 認知科学, Vol. 7, No. 1, pp. 93–106, 2000.
- [水野 22] 水野沙希, 北条伸克, 小橋川哲, 増村亮. 多人数ビデオ会話における非言語情報に基づく次話者予測. 人工知能学会研究会資料 言語・音声理解と対話処理研究会, Vol. 94, No. 0, p. 01, 02 2022.
- [中原 19] 中原和洋. 視線情報を用いた拡散的思考タスクにおけるパフォーマンス推定に関する研究. 人工知能学会全国大会論文集, Vol. JSAI2019, pp. 4Rin115–4Rin115, 2019.
- [中野 14] 中野有紀子, 馬場直哉, 黄宏軒, 林佑樹. 非言語情報に基づく受話者推定機構を用いた多人数会話システム. 人工知能学会論文誌, Vol. 29, No. 1, pp. 69–79, 2014.
- [長谷 14] 長谷川孔明, 中内靖. テレプレゼンスロボットによる無意識的身ぶりの表出が発話交替に与える影響. 日本機械学会論文集, Vol. 80, No. 819, pp. DR0321–DR0321, 2014.
- [飯塚 22] 飯塚陸斗, 川口一画. 多人数ビデオ会議における話者交替のための参与役割の提示手法. Technical Report 31, 筑波大学情報理工学位プログラム, 筑波大学システム情報系, aug 2022.
- [敷田 17] 敷田幹文, ラティカンアルニー. 人数が不均衡な遠隔テレビ会議における弱い光を用いた視線アウェアネス. 情報処理学会論文誌, Vol. 58, No. 1, pp. 166–175, 2017.
- [坊農 04] 坊農真弓, 鈴木紀子, 片桐恭弘. 多人数会話における参与構造分析: インタラクシオン行動から興味対象を抽出する. 認知科学, Vol. 11, No. 3, pp. 214–227, 2004.
- [坊農 08] 坊農真弓. 会話構造理解のための分析単位: F 陣形 (i 連載チュートリアル i 多人数インタラクシオンの分析手法 [第 6 回]). 人工知能, Vol. 23, No. 4, pp. 545–551, 07 2008.

- [櫻井 15] 櫻井翔, 中里直人, 吉田成朗, 鳴海拓志, 谷川智洋, 廣瀬通孝. 表情変形フィードバックによる遠隔協調作業における創造力向上支援 (特集「教育・訓練・支援」). 日本バーチャルリアリティ学会論文誌, Vol. 20, No. 4, pp. 323–332, 2015.

著者論文リスト

本論文に関する論文および発表

- 査読あり国際会議論文
 1. Rikuoto Iitsuka, Ikkaku Kawaguchi, Buntarou Shizuki, Shin Takahashi. Multi-party Video Conferencing system with Gaze Cues Representation for Turn-taking. The 27th International Conference on Collaboration Technologies and Social Computing (CollabTech 2021), Springer, pp. 101–108, August 2021.
- 査読なし国内会議論文
 1. 飯塚陸斗, 川口一画. 多人数ビデオ会における話者交替のための視線提示手法. インタラクシオン 2021, 情報処理学会, 2021年3月, 5 pages.
 2. 飯塚陸斗, 川口一画. 多人数ビデオ会議における話者交替のための参与役割の提示手法. 第199回 HCI 研究会, 情報処理学会, 2022年8月, 8 pages.
 3. 飯塚陸斗, 川口一画. 多人数ビデオ会議における話者交替のための参与役割提示手法 -創造会議におけるシステム評価実験-. 情報処理学会 第85回全国大会, 情報処理学会, 2023年3月, 2 pages (To Appear).

その他論文

- 査読なし国内会議論文
 1. 飯塚陸斗, 川口一画. 記憶性を高める空間音響を用いた通知システムの提案 ～情報の受け取りに適した角度の調査～. 第195回 HCI 研究会, 情報処理学会, 2021年12月, 8 pages.
 2. 飯塚陸斗, 川口一画. 空間音響を用いた仮想的な方向感が付与された音が通知の記憶性に与える効果. インタラクシオン 2022, 情報処理学会, 2022年3月, 6 pages.
 3. 飯塚陸斗, 大山智弘, 杉山将利, 下田康太, 田之頭吾音, 井原圭一, 石橋遼, 川口一画. 動的音高変化による方向提示手法設計のための基礎的調査. インタラクシオン 2023, 情報処理学会, 2023年3月, 5 pages (To Appear).

付録A 実験におけるアンケート

5章にて述べた実験に用いたアンケートを示す。

A.1 条件間に行ったアンケート

各条件後のアンケート

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



*必須

名前を半角ローマ字 (Tanaka Taro) で回答してください *

回答を入力

使用した条件を選択してください (実験アプリ左上の数字) *

0

1

2

次へ

[フォームをクリア](#)

Google フォームでパスワードを送信しないでください。

このフォームは 筑波大学システム情報工学研究科コンピュータサイエンス専攻インタラクティブプログラミング
研究室 内部で作成されました。 [不正行為の報告](#)

Google フォーム



各条件後のアンケート

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



*必須

各条件後のアンケート

以下の項目について自己の評価を10段階で採点してください

アイデアの数はたくさん考えられたか*

1 2 3 4 5 6 7 8 9 10

全く考えられな
かった

とても考えることが
できた

独創的なアイデアを考えることができたか*

1 2 3 4 5 6 7 8 9 10

全く考えられな
かった

とても考えることが
できた

幅広い用途のアイデアを考えることができたか*

1 2 3 4 5 6 7 8 9 10

全く考えられな
かった

とても考えることが
できた

詳細なアイデアを考えることができたか*

1 2 3 4 5 6 7 8 9 10

全く考えられな
かった

とても考えることが
できた



戻る

次へ

フォームをクリア

Google フォームでパスワードを送信しないでください。

このフォームは 筑波大学システム情報工学研究科コンピュータサイエンス専攻インタラクティブプログラミング
研究室 内部で作成されました。 [不正行為の報告](#)

Google フォーム



各条件後のアンケート

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



*必須

各条件後アンケート

以下の項目について5段階で回答してください

今自分が会話に参加していること/していないことを意識することができた *

1 2 3 4 5

とてもあてはまらない とてもあてはまる

自分の話を聞いている参加者がわかった *

1 2 3 4 5

とてもあてはまらない とてもあてはまる

自分がちゃんと聞いていることが相手に伝わっているように感じた *

1 2 3 4 5

とてもあてはまらない とてもあてはまる

自分が話し終えるとき、次に話たさそうな人がわかった *

1 2 3 4 5

とてもあてはまらない とてもあてはまる



自分が話し終わるとき、話したくない人がわかった *

1 2 3 4 5
とてもあてはまらない ○ ○ ○ ○ ○ とてもあてはまる

他の人が話し終わるとき、自分が話すべきかがわかった *

1 2 3 4 5
とてもあてはまらない ○ ○ ○ ○ ○ とてもあてはまる

議論に集中できた *

1 2 3 4 5
とてもあてはまらない ○ ○ ○ ○ ○ とてもあてはまる

議論にストレスを感じた *

1 2 3 4 5
とてもあてはまらない ○ ○ ○ ○ ○ とてもあてはまる

議論に疲労を感じた *

1 2 3 4 5
とてもあてはまらない ○ ○ ○ ○ ○ とてもあてはまる



話者交替にストレスを感じた *

1 2 3 4 5

とてもあてはまらない

とてもあてはまる

戻る

送信

フォームをクリア

Google フォームでパスワードを送信しないでください。

このフォームは 筑波大学システム情報工学研究科コンピュータサイエンス専攻インタラクティブプログラミング
研究室 内部で作成されました。 [不正行為の報告](#)

Google フォーム



A.2 全ての条件後に行ったアンケート

実験後のアンケート

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

条件0、1において、自分が話したい時にシステムはそれを示せたか

1 2 3 4 5

とてもあてはまらない とてもあてはまる

条件0、1において、自分が話したくない時にシステムはそれを示せたか

1 2 3 4 5

とてもあてはまらない とてもあてはまる

システムに対する印象について何か感じたことがあれば記入してください

回答を入力

システムが会話に与えた影響について何か感じたことがあれば記入してください

回答を入力

その他、今回のタスクの中で気づいたことや問題点があれば記入してください

回答を入力



話者交替（話者衝突、沈黙など）について何か感じたことがあれば記入してください

回答を入力

送信

[フォームをクリア](#)

Google フォームでパスワードを送信しないでください。

このフォームは筑波大学システム情報工学研究科コンピュータサイエンス専攻インタラクティブプログラミング研究室 内部で作成されました。 [不正行為の報告](#)

Google フォーム

