

非母語話者のための母語話者との会話を
支援する複数エージェントの研究

筑波大学

図書館情報メディア研究科

2018年3月

唐夢苑

目次

第1章	序論	5
1.1	研究背景	5
1.2	研究目的	5
1.3	構成	5
第2章	関連研究	6
2.1	NNS-NSのコミュニケーションにおける非対称性	6
2.2	第二言語会話支援	6
2.3	多人数コミュニケーション	7
2.4	多人数コミュニケーションへのエージェントの利用	8
2.4.1	会話エージェントの介入方法	8
2.4.2	会話エージェントの多人数会話への支援	9
第3章	NNSヘルプエージェント	11
3.1	対面会話における参与構造と会話モデル	11
3.1.1	2者間の会話モデル	12
3.1.2	3者間の会話モデル	13
3.2	提案	13
3.3	設計	16
3.3.1	開発環境	16
3.3.2	音声検出	17
3.3.3	エージェントの外見	17
3.3.4	エージェントの応答機能	18
3.3.5	介入機構	19
3.3.6	音声合成	20
3.4	実装	21
3.4.1	1体エージェント	21
3.4.2	2体エージェントの連携	23
第4章	実験	24
4.1	1体エージェント	24
4.2	実験方法	24
4.2.1	実験環境	24
4.2.2	実験装置	26
4.2.3	参加者	26
4.2.4	実験条件	26

4.2.5 取得するデータ	28
4.2.6 データの分析方法	30
4.2.7 実験手順	31
4.3 結果と検討	32
4.3.1 発話数	36
4.3.2 話者交替数	36
4.3.3 発話区間	38
4.3.4 発話・沈黙区間の割合	39
4.3.5 質問紙調査	39
4.3.6 検討	41
4.4 2体エージェント	41
4.5 実験方法	41
4.5.1 実験環境	41
4.5.2 実験装置	44
4.5.3 参加者	44
4.5.4 実験条件	44
4.5.5 取得するデータ	45
4.5.6 データの分析方法	45
4.5.7 実験手順	45
4.6 結果と検討	46
4.6.1 発話数	48
4.6.2 話者交替数	49
4.6.3 発話区間	51
4.6.4 発話・沈黙区間の割合	51
4.6.5 質問紙調査	52
4.6.6 検討	54
第5章 結論	55
謝辞	56

図表の目次

図 1. 多人数会話の参与構造	11
図 2. NS-NS の 2 者間の会話	12
図 3. NS-NNS の 2 者間の会話	12
図 4. 3 者間の会話	13

図 5. 提案手法	14
図 6. エージェント介入後の第二言語会話	15
図 7. サンプルエージェントの画面	16
図 8. MMD Agent システムアーキテクチャ	17
図 9. エージェントの外見	18
図 10. 介入タイミングの例	20
図 11. システムの構成	21
図 12. システム処理のプロセス	22
図 13. 2体エージェントの連携	23
図 14. 実験配置 (エージェント発話条件・エージェント非発話条件)	24
図 15. 実験配置 (非エージェント条件)	25
図 16. 実験の様子 (エージェント発話条件・エージェント非発話条件)	25
図 17. 実験の様子 (非エージェント条件)	26
図 18. 実験手順	31
図 19. 発話数	36
図 20. 話者交替数	37
図 21. 話者交替の方向	38
図 22. 発話区間	38
図 23. 発話・沈黙区間の割合	39
図 24. 実験配置 (2体エージェント対面向き条件)	42
図 25. 実験配置 (2体エージェント隣接条件)	42
図 26. 実験の様子 (2体エージェント対面向き条件)	43
図 27. 実験の様子 (2体エージェント隣接条件)	43
図 28. 実験手順	45
図 29. 発話数	49
図 30. 話者交替数	50
図 31. 話者交替の方向	51
図 32. 発話区間	51
図 33. 発話・沈黙区間の割合	52

表 1. エージェントの応答内容	19
表 2. 実験参加順番	27
表 3. 質問項目一覧	29
表 4. 条件ごとに配る質問の項目	30
表 5. S 処理結果一覧(エージェント発話条件)	33
表 6. 処理結果一覧(エージェント発話条件)	34
表 7. 処理結果一覧(エージェント非発話条件)	34
表 8. 処理結果一覧(エージェント非発話条件)	35
表 9. 処理結果一覧(非エージェント条件)	35
表 10. 質問紙調査	40
表 11. 実験参加順番	44
表 12. 処理結果一覧(2体エージェント対面向き条件)	46
表 13. 処理結果一覧(2体エージェント対面向き条件)	47
表 14. 処理結果一覧(2体エージェント隣接条件)	47
表 15. 処理結果一覧(2体エージェント隣接条件)	48
表 16. 質問紙調査	53

第1章 序論

1.1 研究背景

近年、グローバル化と共に、異なる母語の人々間でのコミュニケーションを行う機会が増えている。共通言語を使って非母語話者(NNS)は母語話者(NS)と会話を行う場合、知識や言語能力の不足で、NSの発話に対して応答しにくい、NSとの会話にうまく参加できないなど会話のアンバランスが生じられる。

NSと比べて、NNSの第二言語運用力が弱い立場に置かれているのが事実である。その言語運用能力の低さや文化の差などにより、NNSはNSとの対話において、言いたいことをうまく口に出せない、あまり話題に入れない、あるいはNSは一方的にしゃべるばかりという会話のバランスが崩れる状況は第二言語会話の場面で多く発生している。第二言語会話場面では、NNSとNS両方の協調的なコミュニケーションを実現するために、NNSは多く発話できて、積極的に会話に参加する必要がある。

1.2 研究目的

本研究では第二言語会話におけるNNSが多く発話機会を得られるための会話支援エージェントを提案した。NNSが発話権をうまく取れない場合、エージェントはNNSの意見を伺うような発話をすることで、NNSの次の発言を促すことを目指す。NNS積極的な発話のための介入型エージェントの導入によって、NNSとNSの会話参加のアンバランスが緩和できて、より円滑かつ協調的な第二言語会話が期待される。

1.3 構成

本論文は、本章を含め5つの章で構成されている。第1章では研究背景と研究目的について記述する。第2章では関連研究について紹介する。第3章では提案エージェント及びエージェントの設計と実装について説明する。第4章では提案されたエージェントを使って行った評価実験及び実験結果について述べる。第5章では結論について述べる。

第2章 関連研究

2.1 NNS-NS のコミュニケーションにおける非対称性

第二言語の環境において、NNS は NS とコミュニケーションをする場合、第二言語能力や関連知識が NS より弱いことが知られている。NS は常に発話の優位性を持っていて、発話量が多くて、会話を主導することになっている。

Zuengler の調査[1]では、NNS-NS の会話において、お互いの専門領域以外の話題(例えば、食べ物)について話した場合、発話の言葉数、話題転換の回数など、NNS は NS より少なく、NS は会話への参加度が NNS より高いことが指摘された。同じような知識を持っている、あるいは NNS が比較的乏しい知識を持っている場合、NS はより発話を引き取る傾向が見られることが示唆された。

L Van Lierら[2]は言語能力のレベルがNNS同士間のやり取りにも影響を与えるかどうかを調べた。言語の熟達度が高いNNSは低い者より話題の持ち出し、質問を使って回答を引き出すなど、会話中で積極的な役割を果たしていることを示している。その一方、熟達度が同じようなNNS同士間では会話への参与の違いが見られない、コミュニケーションの対称性が達成できたと述べている。NNSとNS両者の言語能力の違いから、NSが会話をコントロールする可能性が示唆された。

また、W Zhu[3]はNSとNNSが混在しているライティングのグループにおけるメンバー同士のコミュニケーションを調査した。同じグループでのNNSはターンを取ったのが少なかった。特に口頭での討論の中、より少ない言語機能を生み出したと述べている。

そして、NNS-NSのやり取りに存在しているNS主導の非対称性について心理的な角度から調査した研究として、一二三[4]はNNSとの会話におけるNSの言語的処理、内的処理の特徴及びそれらの関連性を検討した。NNSとNSの1対1会話の録音資料やNNSに対する質問紙調査を分析した結果、NSは会話の相手がNNSの場合、「主導的役割の必要性」を感じたことを示した。NNSがNSとの対話をするとき、受動的に会話に参加する立場が示唆された。

しかし、協調的な第二言語会話はNSが一方向的に発話して、NNSが受動的に情報を受けるではなく、NNSが積極的に発言できて、話者交代はバランスを取られるような双方向的なコミュニケーションが必要である。

2.2 第二言語会話支援

NNSの参加している第二言語会話を支援するための研究は多数行われている。音声、文字、機器翻訳など様々な支援手法が検討されている。

音声認識ASRによって、NNSとNSの発話内容をテキストに変換し、またそのテキストをリアルタイムにNNSとNSに共有するツールが作成されて、第二言語会話に応用した研究例が

ある。Gao ら [5] は NS が NNS との多人数コミュニケーションを行うときの共通理解の問題に向け、NNS のバイリンガル辞書の使用状況を NS にも表示するという方法を提案した。またその方法の効果を検証した。NNS がどのように自動化トランスクリプトとバイリンガル辞書を利用するかということ NS は細かく意識させることで、NNS と NS のお互いの共通理解が高くなった結果が示された。第二言語会話に音声認識による自動化トランスクリプトを応用した上、NNS の認知負荷を減少するため、Hautasaari ら [6] は ASR トランスクリプトにあるキーワードが自動的にハイライトされることによって、NNS が音声会議の見逃した部分をうまく追いつけるかどうかを調査した。自動認識されたキーワードをサポートチャンネルとして、NNS はそれを用いて、会話コンテンツをリマインドし、音声会話を素早くレビューできることが示唆された。埜ら [7] は NNS の会話負担を軽減するために、NS が会話中に適宜にキーワードによりキーワード入力することで NNS に共有する手法を提案した。その手法は遠隔環境を模擬したリアルタイムな NNS-NS の第二言語会話に運用する結果、2 話者の相互理解が強化され、お互いに知識の共有を促進できることが表示された。また、音声メディアの処理により NNS を含む第二言語会話を支援する研究例としては、Ye ら [8] は NS 話速により NNS が NS の発話を理解できない問題に対して、話速認識システムを開発した。それを使って、NS の話速をリアルタイムに認識し、速度が速すぎになる際、参加者に知らせるという支援手法を提案した。

NNS-NS の第二言語会話への支援に関する研究は情報通信技術や多様なメディアを利用して、NNS 発話理解上の負担を低減すること、または NNS と NS 間の知識を共有する面にフォーカスして、支援方法を提供しているものが多い。しかし、第二言語会話における発話機会のアンバランスに対して、NNS への支援方法については深く検討されていない。

2.3 多人数コミュニケーション

会話に参加する会話者が会話コミュニケーションとの関わりは常に参与枠組という概念で表れている [9]。参与枠組は発話権から話者を話し手と聞き手に分ける。2 人で構成される一対一の対面会話では、1 人が発話権を持ち、話し手になる場合、もう 1 人が自然的に聞き手になる。しかし、一対一の会話において、話し手が積極的に、多く発言して、聞き手がずっと発話を聞いているままである状況がよく見られている。2 人の会話と異なり、3 人の多人数会話では話し手以外に、聞き手を更に受け手、傍参与者に分類する [9]。多人数会話において、話者の発話権取得に対する能動性によって、発言機会も違う。それは発話の多様化を促すことにつながる。

藤本ら [10] は直接的に相互作用が可能な人数で構成された、協力して課題を遂行するという小集団会話について研究を行っている。小集団会話を円滑に進めるために、役割分担など他の会話者との調整を行う必要性が指摘された [11]。また、その会話展開パターンは初対面

の会話者間で行われる雑談の場面で、他のメンバーの発言を促すような”促進型”が挙げられた。促進型は会話者としてよりも一歩引いたところから会話に参加するが、情報提供もしない発言もしくは先の内容をまとめた発言をするような特徴を持っている。多人数会話において、発言内容の機能性の違いによって会話者は異なった役割を担当し、話者間の関係を調整するとともに、会話の進行や状況に影響を与えることがうかがえる。

会話に参加する人数と規模は会話行動に影響する。少人数会話では、参加者は会話参加の義務が比較的強い、多様な働きをしなければならないと比べて、多人数会話では、話者ひとりの会話への参加義務が軽減される。藤本ら[12]による3者会話と5者会話の調査において、5者会話のほうはより多様な発言パターンが分離できると述べられた。

NNS-NSで構成した会話では、NSはNNSの言語能力の制限のため、常に、会話展開を主導する。それより、会話への参加について、NSはNNSより頻繁に行われ、参加義務が比較的になる。NNSは受動的に発話し、会話に入る傾向が示されている。会話参加者が増えれば、多様な役割を分担する参加者が入るとともに、発話権が再配分できて、NSの参加義務が軽減される。会話参加者における話者間の関係の調整する役割により、NNS-NSの発話参加パターンを改善できると考えられる。

2.4 多人数コミュニケーションへのエージェントの利用

現在、多人数会話において、対話エージェント・ロボットシステムの介入に関する研究は進んでいる。それらの会話システムは会話参加者の言語理解、対話制御、応答文構成という機能を通じて、複数の参加者と積極的なコミュニケーションができるように、協調的な人間・機械が混在している会話の実現のため、重要な役割を果たしている。このように会話システムが多人数会話への活用は会話参加者の発話スタイルを変えられるという研究成果が報告されている。

2.4.1 会話エージェントの介入方法

会話エージェントは複数人の会話において、人のように振る舞い、人と自然的にコミュニケーションができることを重要な要素と考えられ、それを達成するために、多くの研究は会話エージェントの人間同士の会話に適切に参加・参入できる機能に注目している。

杉山ら[13]は多人数会話で複数のユーザの発話に対して応答義務の有無を推定する手法を提案した。提案手法を用いて、ロボットが応答義務のある無しの推定により、関連ユーザの発話に対して適切な応答を返し、会話に参加できる。この研究は会話環境のすべての音を対象として、ロボットから応答義務があるかどうかを推定するための有効な特徴を示した。特に、入力音の識別結果とユーザの動きに関する特徴群が応答義務の推定性能の向上に役立つ。

つと指摘された。ロボットが多人数会話にうまく参加できるための応答行動を正確に判断する面、提案手法の有用性が示唆された。

エージェント介入の際、適切な介入方法が必要とされる。八城ら[14]は話者の優位性に着目し、それに基づくエージェントの適切な介入タイミングや介入内容の決定に関する研究を行っている。提案されたエージェントは二つの部分で構成された。会話者の非言語行動から、優位性を推定して、介入しようとする対象発話者と介入方式を決定できる。また、発話者の発話音声を認識し、キーワード抽出による関連発話内容を生成できる。効果的な介入タイミングや介入内容の決定方法が示された。

エージェントの会話への割り込みタイミングについて他の研究例としては、黄ら[15]はエージェントの適当な介入場面と種類を定義し、それを自動検出するための手法を検討した。Woz実験で収集した音声情報、顔方向、姿勢方向について分析を行って、分析結果に基づいて四つの介入場面を定義した。そして、介入の種類をもとに実験参加者から介入タイミングのデータは機械学習による自動検知を行った結果、タイミングであるかどうかの判定では分類精度77.5%を得たことを示した。特徴量がタイミングの自動検知に寄与することが確認できた。一方、応答文生成について、山内ら[16]は話題誘導の応答文生成手法を研究した。現在の話題と誘導先の話題との関係性を考慮したテンプレートの設計、入力可能な用語の決定を行う手法を取り上げた。話題誘導対話システムの評価では、提案手法の使用により、辞書に登録されている話題間に関しては、話題誘導が高い確率(98%)で実現可能である。雑談会話においても、適当な話題誘導を通じて、応答しにくい話題に窮して、会話がうまく行けない問題を解決できるとうかがえている。

2.4.2 会話エージェントの多人数会話への支援

2.4.1節で記述したような会話介入・調整などの手法を用いて複数人のコミュニケーションの活発化・円滑化を支援している会話エージェントの研究例が多くある。

轡田ら[17]は、多人数会話の活性化を支援し、会話を促進する会話システムの研究を行った。飲み会やコンパの様に多人数の参加者が場に存在する状況において、会話グループの構成は時間の経過と共に、変化していく特徴を明らかにした。その上、多人数会話での会話の継続が困難になる問題に対して、会話グループの参加者それぞれの音声データを取得、発話有無の判断を行うことで、会話グループの状況をリアルタイムに認識し、会話に参加していない者を検出する。また、その参加者に何らかの話題を提示する会話支援システムを提案した。

複数人の会話ににおける会話参加の優位性が低い者の発話を促すことについて、高瀬ら[18]の研究では、会話中の参加者の優位性や参与役割を推定して、適切な注目行動を行うことで適切な視線フィードバックを表出できる。それに加えて、発話権がある良いタイミングを判断してそこで参加者の意見を求めるなど、議論を発展させるための発話を行う対話ロボ

ットを作成した。評価実験結果として、提案された対話ロボットは優位性や発話権を考慮した会話介入を行って、会話への参加度が低い者がより多くの発話を促すことが可能であることが示された。

また、多人数会話において発話機会の不均衡の問題に対して、秋葉ら[19]は会話ロボットから適切な手続きを経て、発話機会を与えるための計算モデルを提案した。その研究では、4人会話を対象として、会話の場で発話を主導している者の状況を考慮した上、段階的に主導権を取って、対象となる参加者に発話権を渡すことができるロボットをデザインした。提案手法を通じて「置いてけぼり」状態の参加者は場の一体感が生まれやすいということが指摘された。

エージェントが積極的に会話参与に重点を置いている多人数コミュニケーションでは複数の参加者それぞれはグループ構成員の機能的役割を持っている[20]。エージェントとかロボットなどのシステムは多人数会話の一員として、会話に介入し、適切な機能を働かせて、会話をうまく調整し、参加者間の関係を協調している研究例が多く取り上げられる。特に、エージェントがグループ調整役割を演じて、多人数会話における発話のアンバランスを解決する。第二言語での会話の場において、エージェントからの会話の状況の調整について深く検討されていない。

本研究では多人数グループにおける会話を協調できる役割の特徴を利用し、NNS-NSの1対1の第二言語会話に会話の発話機会を調整できるエージェントを導入し、適切なタイミングを検知した上、NNSへ適切な発話内容を与えることで、NNSにより多くの発話機会を提供して、NNSを支援する。エージェントがNNSへ適当に発話権を付与することによって、円滑かつバランスの良い第二言語会話の実現を期待している。

第3章 NNS ヘルプエージェント

第二言語会話における NNS がうまく発話権を取れない、一方的に NS の発話を聞く問題を解決するため、多人数会話では、会話の進行や発話権を調整できる参加者の役割をシミュレートして、NNS へ適当に発話権を渡すことができるエージェントを提案し、NNS を含む 1 対 1 の第二言語会話に導入する。

3.1 対面会話における参与構造と会話モデル

社会的集団である会話に参加する者は他者の存在を意識し、他者と関わって自分の位置付けを意識する[21]。Goffman は会話における参与構造の存在を指摘している。中核的役割から、話し手(speaker)、受け手(addressee)、傍参与者(side participant)、傍観者(bystander)、盗み聞き者(eavesdropper)のように分類されている。Clark は Coffman が考えている参与構造をまとめた(図 1)。

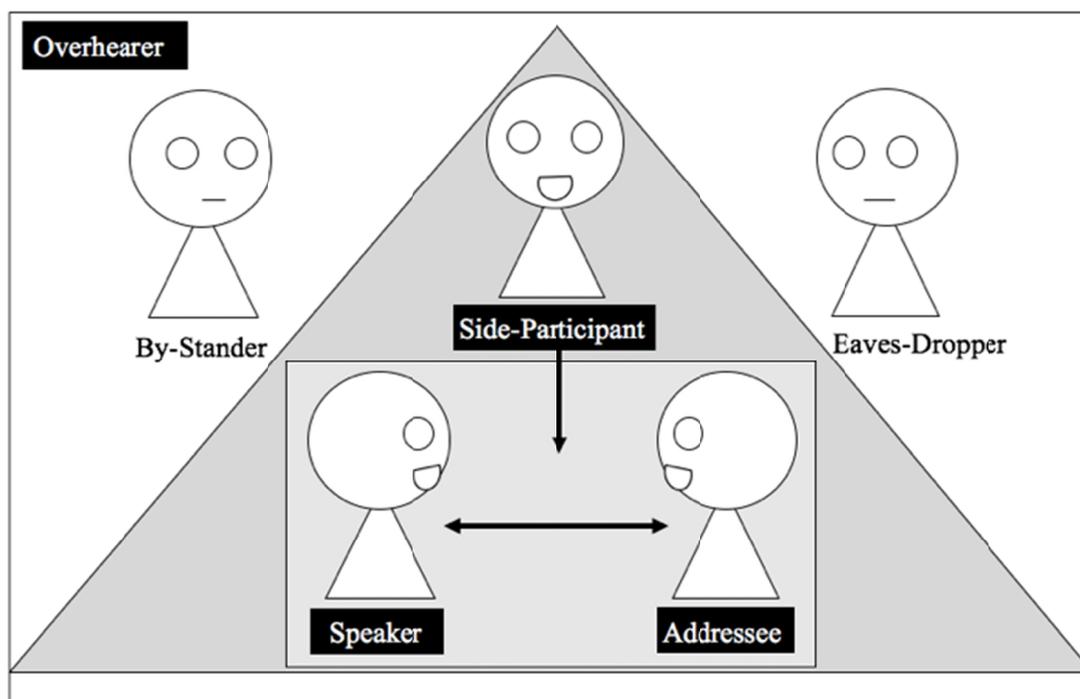


図 1. 多人数会話の参与構造

対面会話において、発話している者は話し手であり、3人以上の会話の場合、聞き手に対して、発話を受ける人が受け手と呼ばれ、受けない人が傍参与者とされている。話し手、受け手、傍参与者は多人数会話の中核である。会話は少なくとも話し手と受け手の間で話者交代が行われるが、残った者は会話を傍観しているまたは会話の中核に参加することがある。

このように、話し手と受け手以外の第三者などは既存の参与者間の関係を調整できて、会話の成り行きをコントロールすることができる。

3.1.1 2者間の会話モデル

2者間で会話を行う場合では、2人の参加者は話し手と聞き手の間に変換する。2者間で発話権の所在や文脈がお互いに了解しながら会話が進む[19]。NS-NSの1対1の会話で、お互いに同じ言語レベルや文化的な背景を有しているため、二人の会話参加者は会話の中、交互的に発話を交わして比較的均衡に発話権を取られると考えられる(図2)。しかし、第二言語で会話する場合、NNSが言語的な劣勢を持っていて、NNSとNSは会話が進みながらお互いに共通理解基盤を作るために多様な会話スキルやストラテジーを使用しているが、会話の展開のため、NSはより多く発話権を持って、会話の主導的な役割を果たしている(図3)。

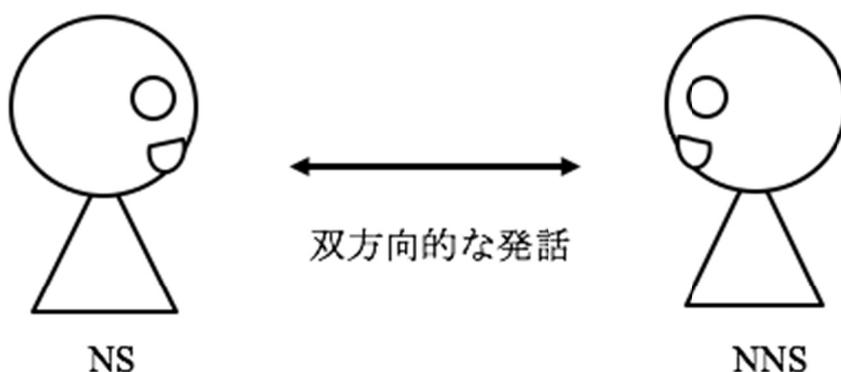


図2. NS-NSの2者間の会話



図3. NS-NNSの2者間の会話

3.1.2 3者間の会話モデル

3者間の会話では、話し手と聞き手は会話の中核を構成しているが、第三者である傍観者は積極的に会話に参加することがある。このように、傍観者は今の参加者間の関係を変更し、発話権を奪って話し手になり、会話の進行をコントロールする可能性がある。

発話権がうまく配られない NNS-NS の二者間の会話においては、その発話が不均衡している状況を調整できる役割を担う第三者が会話に加入すれば、NNS と NS の参加関係を積極的に改善すると考えられる(図4)。

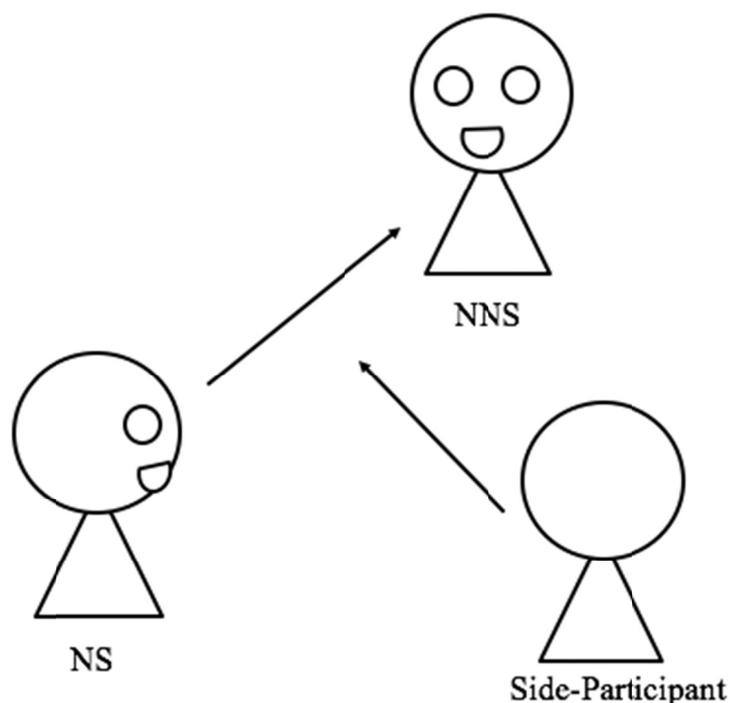


図 4.3 者間の会話

本研究では NNS ヘルプエージェントを NNS-NS 一対一の第二言語会話に導入し、3人会話の場をシミュレートする。その場でエージェントを発話権調整の役割を担当させて、会話中に NNS へ適当に意見などを伺うことによって、発話権を NNS に渡すという手法を取り上げる。このように、話者交代を行って、NS から一方的に発言し、会話を主導するような不均衡な会話状態を改善し、NNS が多くの発話機会をもらうことによって、より協調的な第二言語会話の実現に役立つことを望む。

3.2 提案

NNS ヘルプコミュニケーションエージェントを第二言語での会話に導入させる。エージェントは2話者の音声を継続に検出していて、会話を介するかどうかとリアルタイムに判断す

る. NS と NNS 両方が一定の沈黙に入る場合, 介入すると判断して, NNS へ伺いの内容を決める. また, 決定された内容を音声合成技術により, NNS 宛てに発話を返して, 次の発話権を NNS に移動させるという手法を提案する. 図 5 は提案手法の流れを示している.

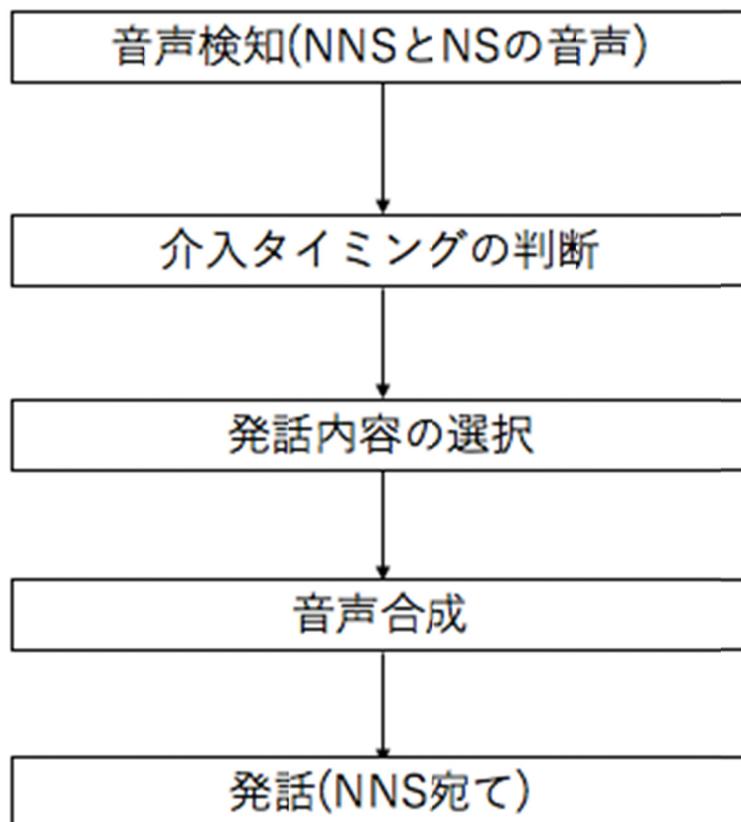


図 5. 提案手法

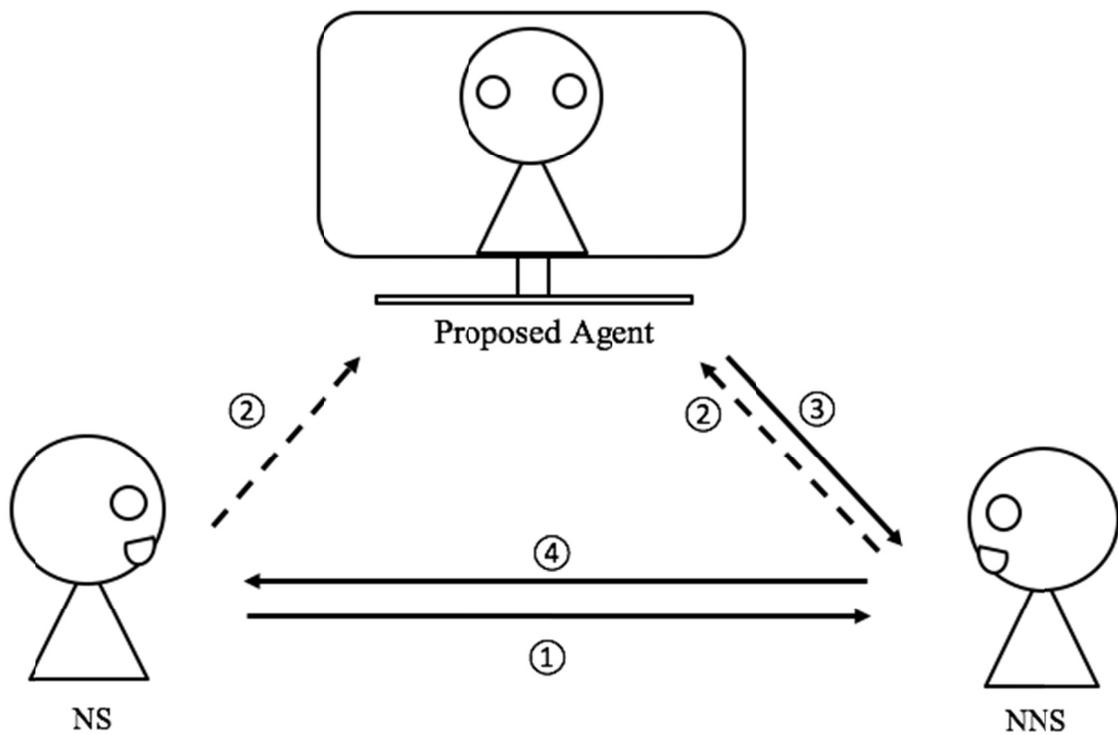


図 6. エージェント介入後の第二言語会話

エージェント介入後の第二言語会話は図 6 のように表す. ①NS が NNS 向け発言する②エージェントが NS・NNS の発話を検知する, 両方が一定の沈黙に陥る場合, NNS に伺うような自己の発言を選択する. ③エージェントから NNS 宛て話を伺う. ④NNS が話者権をもらい, NS へ発言する.

3.3 設計

3.3.1 開発環境

名古屋工業大学国際音声技術研究所によって開発されたオープンソースの音声インタラクションシステムを構築するためのツールキット(MMD Agent)を利用した[22]. MMD Agent は 3DCG, 音声合成, 音声認識, 対話制御などの機能を1つにまとめたコンテンツを持っている. また, MMDA という独自の拡張子を持っている. MMD Agent はすべての素材をそのまま使用したり, または既存の他のシステムの1つ以上のモジュールを利用したりすることができるようにする. この技術は, 自然言語の相互作用, 非言語的な行動を工夫している. 他には新しい技術を構築するための広く受け入れられたプラットフォームを提供できる. 本研究は人間と似ている三次元のキャラクタを必要とし, 人と言語や非言語的なインタラクションを行われる基本機能を有しており, また複数のエージェントを協調制御できるような拡張機能を実現できる開発環境が提供されると考えた上, このツールキットを利用した. MMD Agent のサンプルエージェントとシステムのアーキテクチャは次の図7と図8で示す.

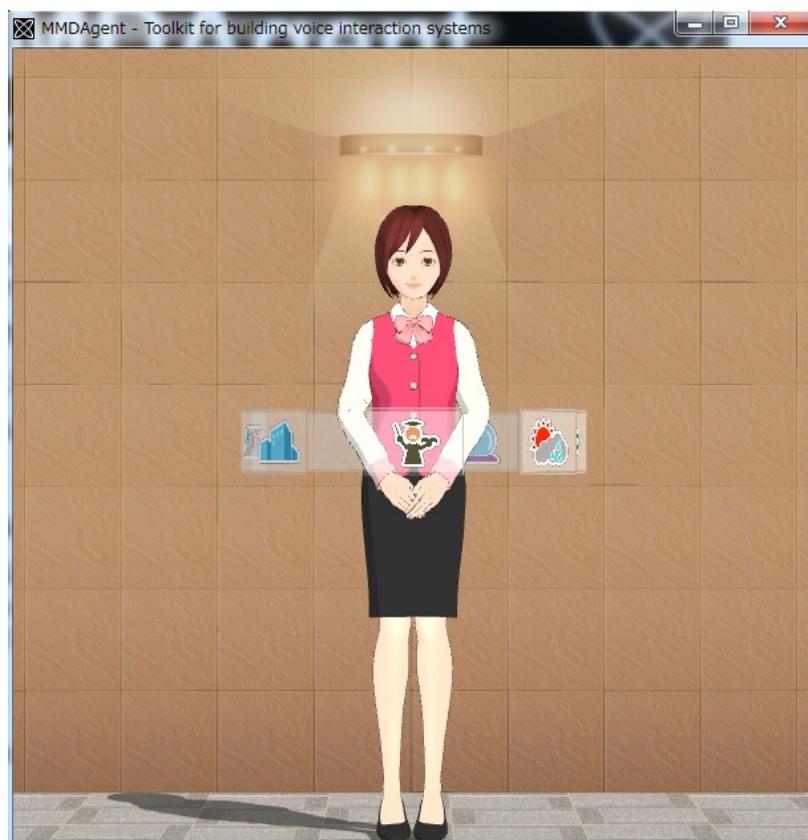


図7. サンプルエージェントの画面

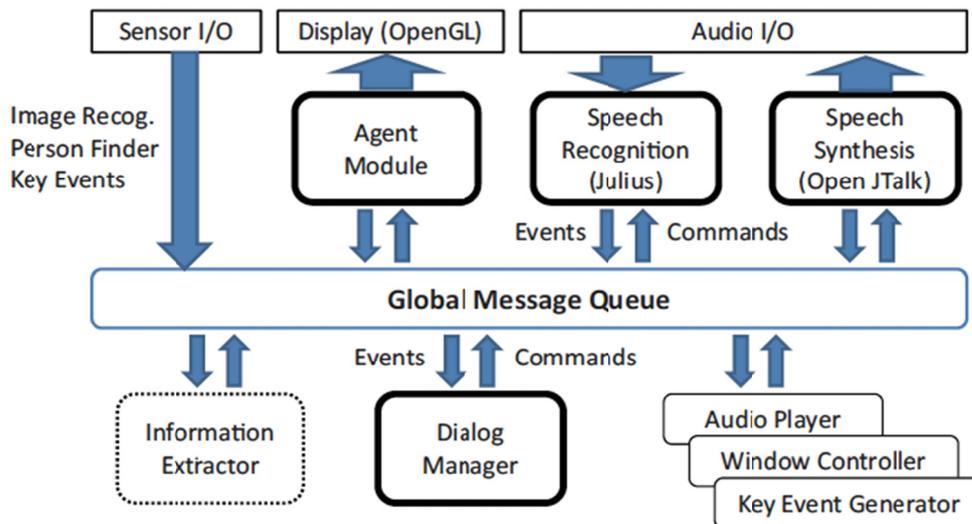


図 8. MMD Agent システムアーキテクチャ [23]

3.3.2 音声検出

マイクから入力された音響データをリアルタイムに取得するために、Python 環境で PyAudio[24] というオーディオ I/O ライブラリを利用した。

音響データは 16bit の PCM データ (Pulse Code Modulation: パルス符号変調) の形で取る。普通の会話の音量は 60db [25] 前後であり、振幅値がおよそ 1500 に当たるため、発話の閾値を 1500 にする。NNS と NS の音声の大きさは 1500 を超えたら、発話と判断する。

3.3.3 エージェントの外見

エージェントは第二言語会話の場面に導入され、会話参加するため、NNS や NS と一緒に会話する人物を考え、等身大の仮想キャラクターが必要である。そこで、MMD Agent が提供するサンプルエージェントの素材をベースに、塩原[26] が作っているエージェントモデルを利用して、一部の調整と変更を施した。調整のためのプラットフォームは 3DCG オブジェクトを作るための MikuMikuDance という 3DCG 作成ソフトウェアを使用した。エージェントは会話の場面に合わせて、座る姿勢にした。(図 9)。



図9. エージェントの外見

3.3.4 エージェントの応答機能

エージェントはNNSが発話権を取れない場合、エージェントは発言して発話権を奪うことで、傍参与者から話し手に変換し、会話に介入する。このようにNNSを次話者として選択するために、NNSに何うような発話内容を生成する必要がある。

サックスら[27]は現在の話し手が次の話し手を選ぶための技法は次の4種類を指摘した。

- 1) 発話のやり取りを分析するための最小単位である隣接ペアにおいて、「質問-応答」というパターンがある。話者選択の際、隣接ペアの第一成分とする質問部を利用できる。特定の宛先を持つ質問で次の話し手を選ぶ。
- 2) 直前の発話の一部分を質問音調で繰り返すあるいは「何?」、「誰?」などの一語質問をする。
- 3) 「でしょ?」、「そう思わない?」、「だよね」などの付加質問をつける。付加質問自体の完了点において、特定の次の話し手の順番が始まる。
- 4) 社会的に何者であるかを利用することで、発話の宛先を特定する。例えば:「夫」、「教員」、「父親」などである。

この4種の選択技法においては、1)は質問を投げることで特定の対象を選択でき、発話権を渡す形である。エージェントから質問を出すことでNNSが特定の対象として次発話者と選ばれると考える。また、2)の選択技法は前話者の発話に基づいて、質問を投げる形として、前話者を次の話し手として選択する方法である。この方法を用いて次発話者を選ぶ際、次の話し手は前話者次第である。そうすると、NSが話した後、エージェントはこの技法を使っても発話権をNSからNNSに渡すことができないと考え、本研究では2)を対象外にした。そして、3)の場合、付加質問を利用して、発話権を特定の対象者に移動させる。エージェントは付加質問をつける発話を発せられて、NNSに発話権を渡すことができると考えられる。

しかし、本システムを使って実験を行う参加者(NS, NNS)は全部初対面の学生であり、また第二言語での会話場面を考慮した上、4)の技法を使用できる場面がない。このように、本システムは1)と3)の方法のみを利用する。エージェント応答内容のデザインについては、NNSに発話権を渡すので、NNSを特定の宛先として、NNSの呼称、「〇〇さん」をつけて発話内容の先頭にする。そして、1)と3)の方法を利用する際、会話中でエージェントが自然に会話に介入させるために、話者の発話内容に関わらず、何の場面でも通用できるかつ一番簡単な質問や付加質問のパターンを考慮した。このように、1)の選択技法を用いた応答内容は「質問-応答」という隣接ペアの質問部は意見を伺う形にして、NNSが意見を話すように次話者として選択される。詳しくは「どう思いますか?」、「どうでしょうか?」のような質問文が2つを用いた。3)を用いた応答内容は、「そう思わない?」、「だよね」という二つの付加質問を発話の最後尾にする。設けられた応答は「〇〇さん、そう思わないの?」と「〇〇さんもそう思いますよね」2つである。トータルは4つの質問を用いた。次の表1は設定した4つの応答内容を示す。

表 1. エージェントの応答内容

番号	応答内容
1	〇〇さん、どう思いますか?
2	〇〇さん、どうでしょうか?
3	〇〇さん、そう思わないの?
4	〇〇さんもそう思いますよね

エージェントは毎回会話に介入すると判断すると、4種類の応答内容から同じ確率で1つを選んで出力する。

3.3.5 介入機構

NSとの会話では、NNSが話していて、発話権がNNS側にある場合、エージェントが何もしない。逆にNNSが発話権をうまく取れない場合、エージェントが介入するという設計指針を定める。エージェントの介入タイミングは会話の順番交代のところに着目する。

サククスら[27]は順番交代システムについて、順番の組み立てを支配する基本的な規則群は順番の移動を調整する。規則群は次の3種類が指摘される。

- 1) 現在の話し手が次の話し手を選択するための技法を用いる形で、選択された者が次の発言順番を取る。
- 2) 現在の話し手が次の話し手を選択するための技法を使用せず、次の発話者になろうとする者は自己選択という方法を使う形で、最初に話し始めた者が次の発話権利を得る。
- 3) 現在の話し手が次の話し手を選択するための技法を使用せず、他の者は自己選択もし

ない場合、現在の話し手は話し続けても良い。

NNS-NS の会話の場面でも、以上の順番移行の規則群が使われている。このように、NS は発話の主導権を持っていて発話した後、順番の移行に適切な最初な場所に至った時、規則群により、①NS から NNS を次の話し手として選択する。②NS が NNS を選択しない、NNS が自己選択によって次の発話権を取る。③NS は次の話し手を選択しない、NNS も自己選択を行わない、NS はまた話し続けるというように順番交代を 3 種にも分類できる。①と②の場合、NNS が次の話し手になり、発話権を取れる。そのため、エージェントの介入が適切ではない。③の場合に、NNS が次の発話権を失うので、エージェントが介入する適切な場面と考え、エージェント介入を決定する。③のように、順番移行に行って、NS と NNS も次話者を選択しなければ、沈黙区間に進む。2 秒間の無音区間が続いて、前話者また話しても別のターンと見られるということで[28]、会話中が無音区間 2 秒を超えたら、エージェントが会話に介入すると判断するように設定した。エージェント会話介入の例は図 10 で示す。

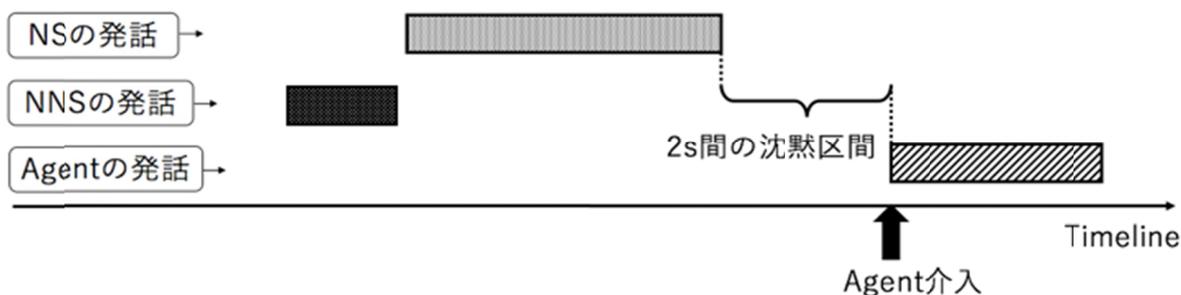


図 10. 介入タイミングの例

3.3.6 音声合成

MMD Agent は日本語のテキストを読み上げる TTS システム Open JTalk を使用している。Open JTalk[29]はキャラクタ「mei」という女性話者モデルの声を提供している。Open JTalk は安定した音質を合成できる。Open JTalk を組み込んでいる MMD Agent は人間の声と近い音声を発することができる。本システムは「mei_normal」の音声を使っている。

3.4 実装

3.4.1 1体エージェント

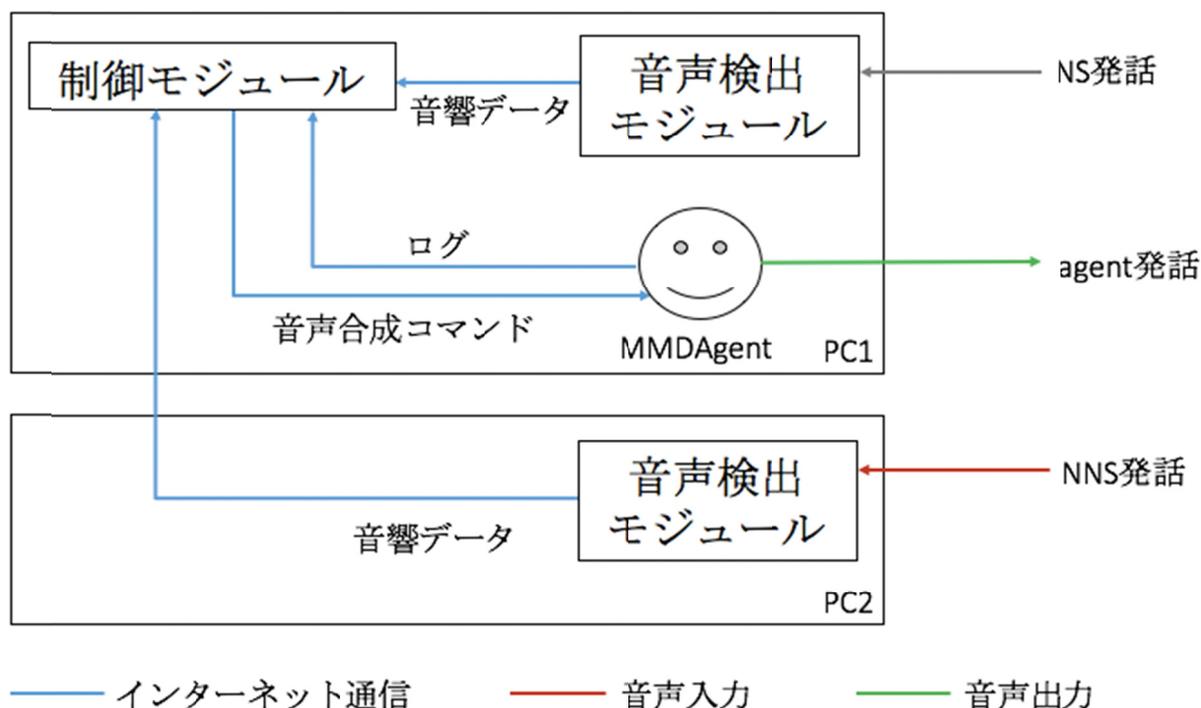


図 11. システムの構成

図 11 で表示しているように、NS と NNS それぞれの音声を単独に検出するため、2 台のパソコンに音声検出のモジュールを載せた。それぞれは NS と NNS 発話の音響データを取得する。1つのパソコンに MMD Agent を搭載した。そのエージェントは図の PC1 に搭載する MMD Agent で示している。そして、このパソコン(PC1)に制御モジュールを持って、2つパソコン間の制御を行う。また、制御モジュールによって、会話介入を判断し、音声合成を通じて、発話することで、会話に参加する。もう1つのパソコンは図の PC2 で表している。このパソコンは NNS の音声検出のために使用する。

特に、MMD Agent のコントロールは、MMD Agent にソケット通信プラグインを追加することによって、制御モジュールと MMD Agent 本体との通信を実現できた。制御モジュールは python に基づいたプログラムである。ソケット通信によって、PC 二台の音響データやエージェントのログをリアルタイムに取得でき、コマンドを送信することでエージェントを制御することもできる。制御モジュールは、音響データと MMD Agent のログに応じて、NS と NNS とエージェントの状態を把握し、それに依りて介入タイミングを決定する。介入する際に、python の random. choice 関数により、4 種類の応答内容から一つを発話内容として選択す

る. MMD Agent に音声合成のコマンドを送ることによって, エージェントを発話させる. 全システム処理の流れは次の図 12 で表す.

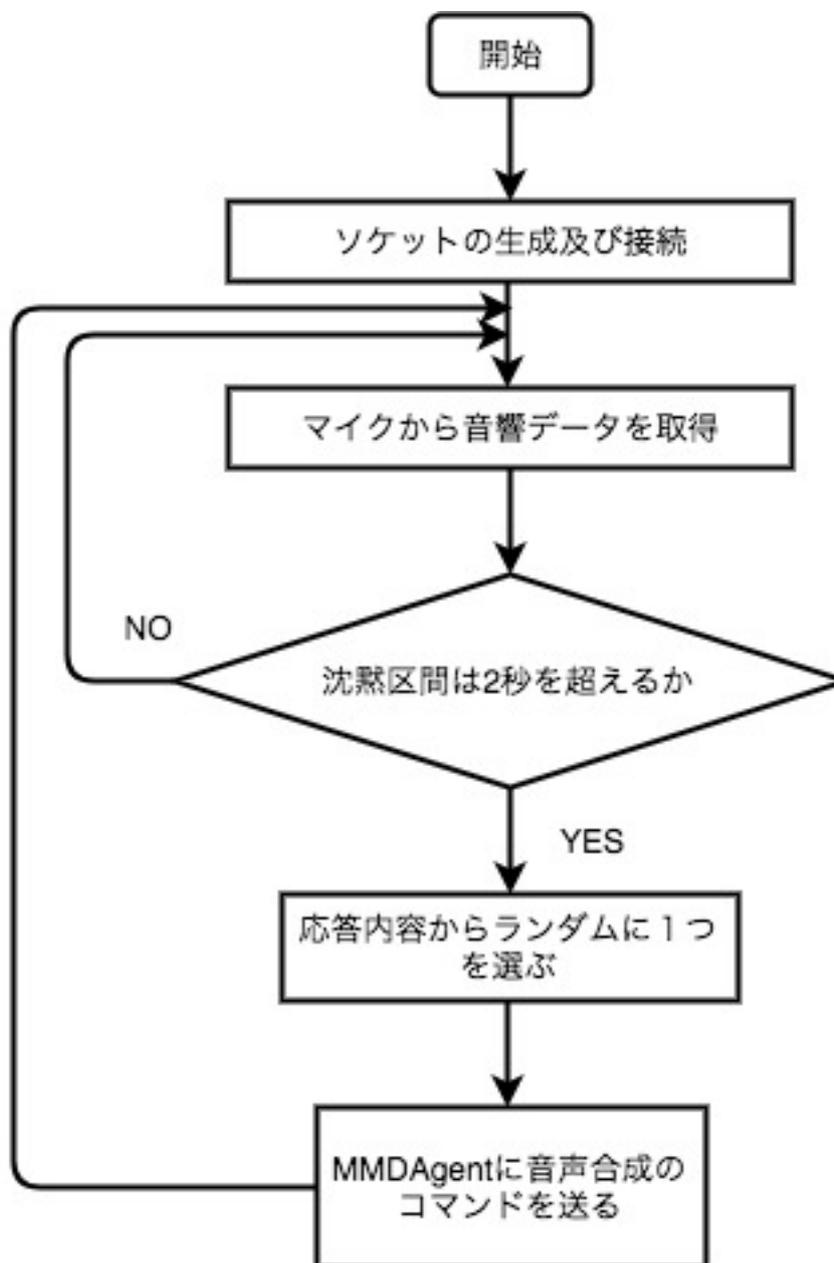
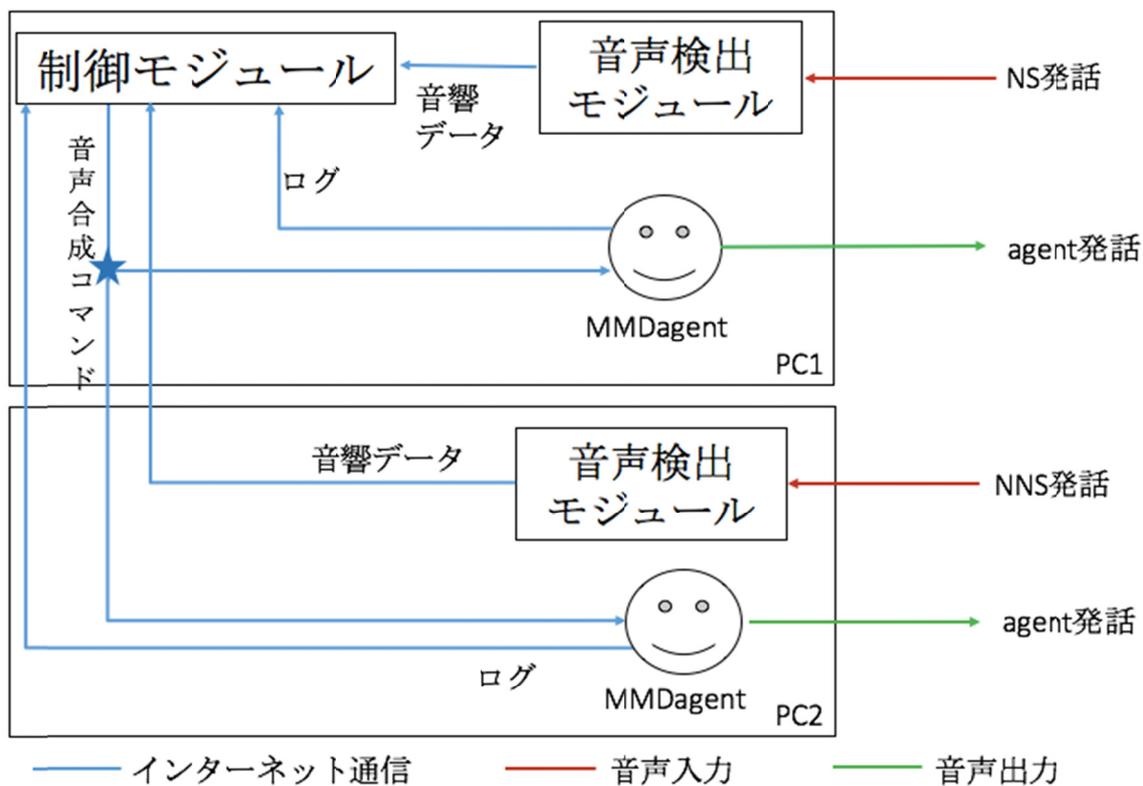


図 12. システム処理のプロセス

3.4.2 2体エージェントの連携



★ 2体のagentからランダムに1体を選んで音声コマンドを送る

図 13. 2体エージェントの連携

2体エージェントを連携する場合、2台のパソコンは(PC1, PC2)共に MMD Agent を搭載した(図 13)。PC1 の制御モジュールは、2つのエージェントをコントロールする。エージェントが会話に介入すると判断したら、制御モジュールによって 2つのエージェントからランダムに1つを選択し、また、選んだ MMD Agent に音声合成のコマンドを送って、そのエージェントを発話させる。

第4章 実験

4.1 1体エージェント

本実験の目的はNNSに発話権を渡すという機能を有するエージェントを第二言語会話に介入することで、エージェントがいない会話と比べて、NNSがより多く発話機会を得るかどうかを評価する。

4.2 実験方法

4.2.1 実験環境

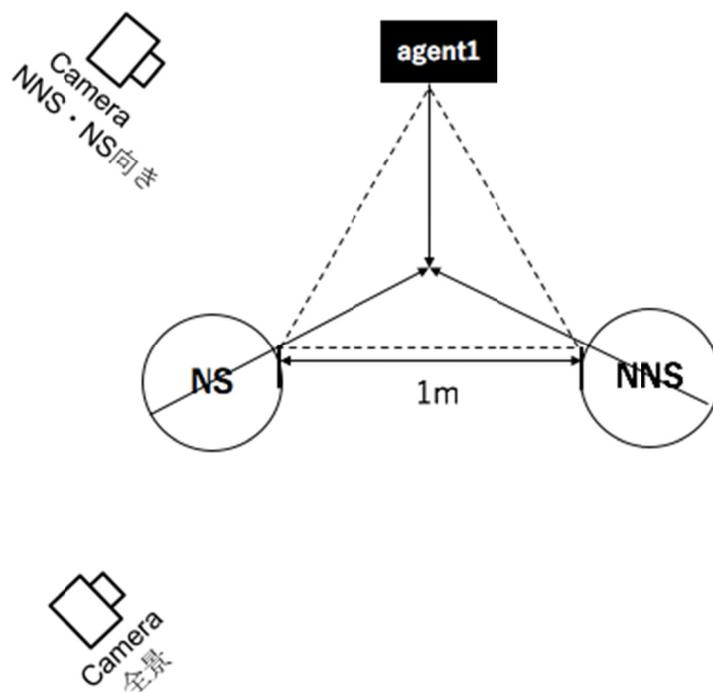


図 14. 実験配置 (エージェント発話条件・エージェント非発話条件)

実験は静かで1対1会話の環境を考慮して、会話用の椅子を用意して、研究室で行った。実験の環境騒音は40dbである。また、本実験では被験者間の対話距離は自然な対話距離として、1mである。エージェントがエージェントを含む条件では、NNS、NS及びエージェント3者間の距離は等間隔にして、3者が三角形の中心に向かう(図14)。非エージェント条件では、NSとNNSは真正面に向かう(図15)。エージェントがいる条件ではカメラ2台が設置される。1台のカメラでNNS、NS及びエージェントを含む全景を撮影し、1台のカメラでNNS

とNSは顔や動きがはっきり見えるように正面から上身を撮影する。非エージェント条件では、エージェントがないため、1台のカメラを用い、NNSとNSの様子を撮影する。

Camera
NNS・NS向き

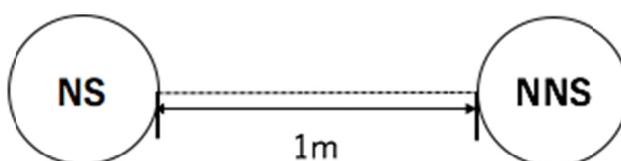


図 15. 実験配置 (非エージェント条件)

そして、各話者の音声を入力し、マイクは話者の口にほど近いところに装着してもらう。会話の様子はエージェントを含む条件が図 16 に示され、非エージェント条件が図 17 に示される。

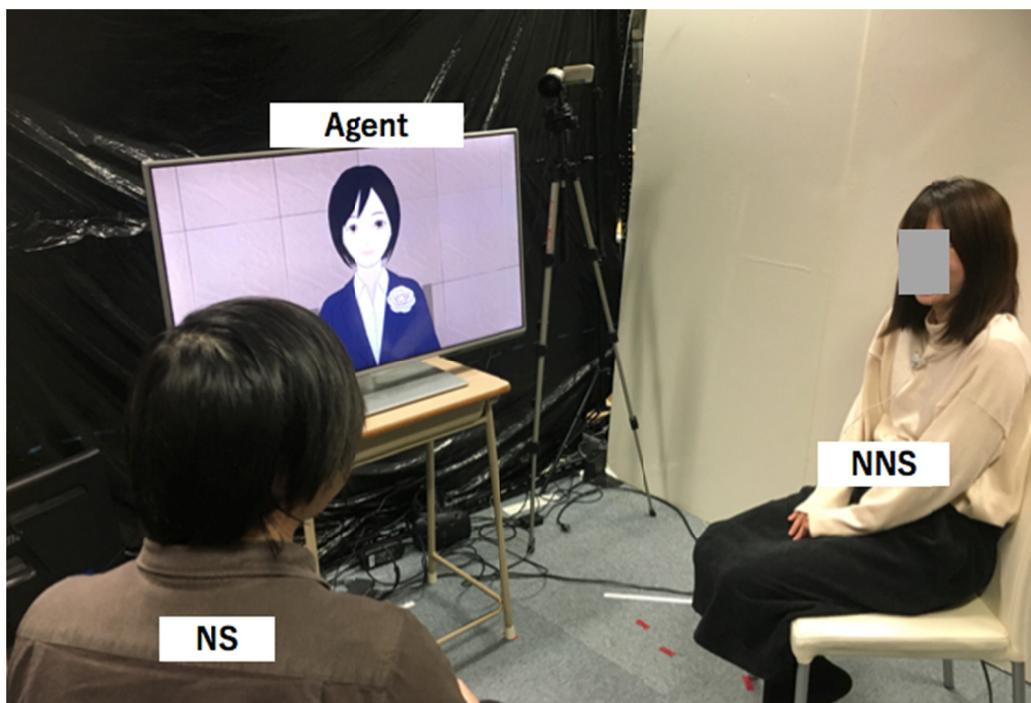


図 16. 実験の様子 (エージェント発話条件・エージェント非発話条件)

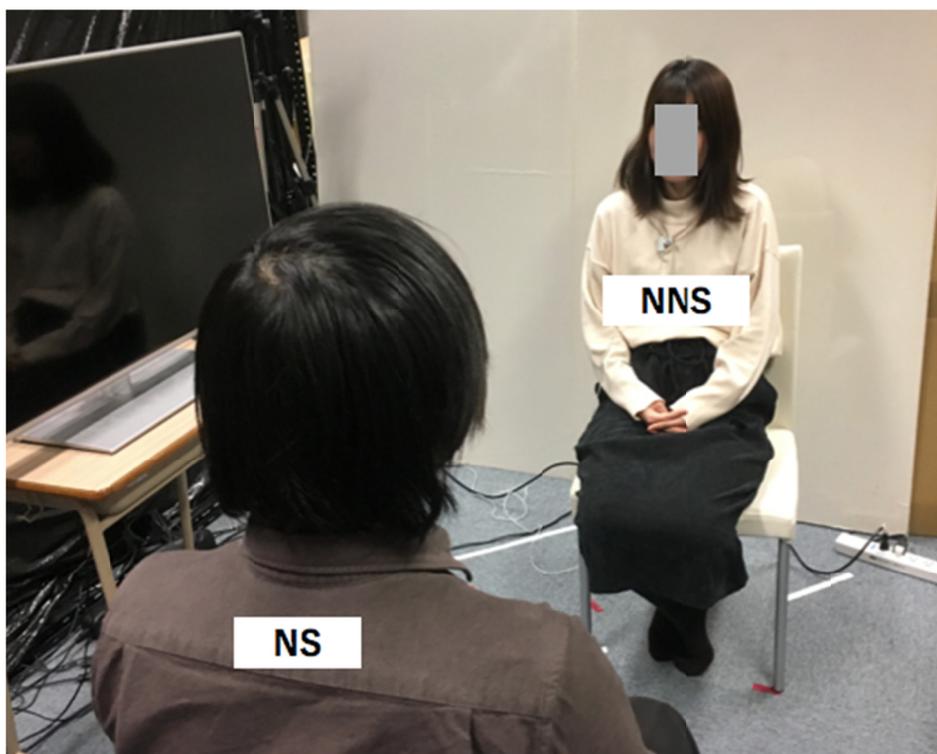


図 17. 実験の様子(非エージェント条件)

4.2.2 実験装置

話者音声の検知と認識は2台のパソコンによって行われる。そのうち、1台はエージェントのコントロールや制御を実現する。音声の検出は単一指向マイクを2個使用する。周波数は80-15,000Hz、感度は-38dBである。NNSとNSの音声はそれぞれのマイクを通じて入力できる。エージェントの表示はパナソニック社の42型(インチ)デジタルハイビジョン液晶テレビを使う。このモニターはサイズが幅957mm×高さ×619mm×奥行217mmであり、画素数が1920×1080である。このモニターを使って等身大のエージェントを映ることができる。撮影機械について、全景及び各話者とエージェントを撮るため、0.7倍のワイドコンバージョンレンズを取り付けたSANYO Xacti DMX-HD2000を2台使った。解像度は1920×1080である。

4.2.3 参加者

初対面の日本語母語話者NS、および日本語非母語話者NNS、各18名からなる18組、計36名である。性別からの影響を小さくするよう[30]各ペアは同性別にする。

4.2.4 実験条件

エージェントからのNNS-NS会話への影響を調べ、被験者内計画を利用する。本研究では1

要因3水準の被検者内計画を立てる。エージェントを会話に影響を及ぼす1要因として、会話中でエージェント発言の有無及びエージェントがいないという三つの条件はこの要因を構成する水準としている。

実験の条件は次の3条件とした。

① エージェント発話条件

エージェント会話を介して、会話の中、NNSに話を伺うことで、発話権を調整する。

② エージェント非発話条件

エージェントはNNS-NSの会話に参加しても、発話しない。

③ 非エージェント条件

エージェントなしで参加者は対面の会話を行う。

順序効果を相殺するため、参加者と3条件を組み合わせて実験の実施順序は以下のように示す(表2)。

表2. 実験参加順番

参加者(組)	実験条件の順番		
1	①	②	③
2	①	③	②
3	②	①	③
4	②	③	①
5	③	①	②
6	③	②	①
7	①	②	③
8	①	③	②
9	②	①	③
10	②	③	①
11	③	①	②
12	③	②	①
13	①	②	③
14	①	③	②
15	②	①	③
16	②	③	①
17	③	①	②
18	③	②	①

4.2.5 取得するデータ

NS と NNS 及びエージェントが行う自由会話の様子を記録した映像および音声データを得る。また、会話後に NNS と NS に対して質問紙による調査をする。

質問紙は異文化コミュニケーションや会話分析で使われているコミュニケーションの自然さ、発話しやすさなどコミュニケーションの状況を評価する項目及び会話全体への印象を表示する項目[31][32][33][34]を利用する。また、システムの有用性を評価するため、エージェント介入の適切さ、エージェントから NNS の発話に役立つかどうかといった項目を設けた。

質問紙は各条件で話者別に用意される。エージェント発話条件では、項目 1 から 6 まで計 6 問はコミュニケーションの状況に関する設問である。それぞれは項目 1「自分の言いたい事が話しやすかった」、項目 2「相手と自然にコミュニケーションできた」、項目 3「相手とまた話したいと感じた」、項目 4「言いたいことが遠慮せずに言えた」、項目 5「頻繁に発言することができる」、項目 6「対等的な会話ができる」である。

会話への満足度に関して項目 7 から 9 まで 3 問を設けられる。項目 7「対等的な会話ができる」、項目 8「会話が楽しかった」、項目 9「会話に参加している感じがした」項目 10 から 14 までの 5 問からなる。

システムの有用性について、項目 10 から 16 までを設ける。項目 10「エージェントは会話に参加している感じがした」、項目 11「エージェントがいるのは自然だったと思う」、項目 12「エージェントは相手の会話への参加に役立った」、項目 13「エージェントは自然に会話に介入できた」、項目 14「エージェントは相手の発言を促すことができたと感じた」、項目 15「エージェント介入によって相手が多く発言できた」、項目 16「エージェントの発話タイミングが適切だと感じた」。ただし、NNS の質問紙で、項目 12、項目 14 及び項目 15 は「エージェントは自分の会話への参加に役立った」、「エージェントは自分の発言を促すことができたと感じた」、「エージェント介入によって多く発言できた」に変更する。

また、話者の個人面におけるストレスに関して、5 問を使う。項目 17 から 21 までそれぞれは項目 17「相手との会話は不快であった」、項目 18「相手との会話が退屈だと感じた」、項目 19「相手と会話していて疲れた」、項目 20「自分はゆっくりできた」、項目 21「自分は気楽にできた」である。このように、設けられた設問はトータル 21 問を有する。

エージェント非発話条件はシステムの有用性について、項目 1 と項目 2 しか使わない。他の設問はエージェント発話条件と同じである。合計 16 問である。

非エージェント介入条件はシステムに関する 7 問を使用しなく、残った 14 問だけを設けられる。評定尺度は 7 段階とする。7 段階は「全く当てはまらない」、「当てはまらない」、「あまり当てはまらない」、「どちらともいえない」、「ややあてはまる」、「当てはまる」と「よく当てはまる」になる。

表 3 はカテゴリーを分けて整理された質問項目一覧表である。条件ごとに配る質問の項目

は表 4 で表示する.

表 3. 質問項目一覧

カテゴリー	番号	質問項目
自分のコミュニケーション	①	自分の言いたい事が話しやすかった。
	②	相手と自然にコミュニケーションできた。
	③	相手とまた話したいと感じた。
	④	言いたいことがうまく言えた。
	⑤	頻繁に発言することができる。
	⑥	対等な会話ができる
会話満足度	①	会話が楽しかった。
	②	会話に参加している感じがした。
	③	会話が盛り上がった。
エージェントへの評価について	①	エージェントは会話に参加している感じがした。
	②	エージェントがいるのは自然だったと思う。
	③	エージェントは自分/相手の会話への参加に役立った。
	④	エージェントは自然に会話に介入できた。
	⑤	エージェントは自分/相手の発言を促すことができたと感じた。
	⑥	エージェント介入によって(相手が)多く発言できた。
	⑦	エージェントの発話タイミングは適切と感じた。
ストレス	①	相手との会話は不快であった。
	②	相手との会話が退屈だと感じた。
	③	相手と会話していて疲れた。
	④	自分はゆっくりできた。
	⑤	自分は気楽にできた

表 4. 条件ごとに配る質問の項目

質問項目	NNS/NS		
	エージェント発話条件	エージェント非発話条件	非エージェント条件
自分のコミュニケーション	①～⑥	①～⑥	①～⑥
会話満足度	①～③	①～③	①～③
エージェントへの評価について	①～⑦	①～②	
ストレス	①～⑤	①～⑤	①～⑤

4.2.6 データの分析方法

初対面で行った自由会話で被験者は最初にある程度自己紹介を行うため、最初の1分間の会話を分析しない、7分の会話を分析対象とする。収集された音声データはアノテーションツール ELAN[35]を使って、NS、NNS 及びエージェントの発話をラベリングする。発話の多寡を測るため、発話数や発話時間という指標を用いる。会話中で話者が得る発話機会を調べて、話者交替数を数える。そして、発話権が移動する方向を詳しく考察するため、話者が誰から発話権を受けるかに分けて、話者交代数を計算する。また、沈黙区間が長く続ければ、会話がうまく進めないと考えられるので、会話の様子を知るため、全会話の沈黙区間を測り、また、話者沈黙の割合及び沈黙区間の分布状況を調べる。

1) 発話数

間終止単位 (IPU) [36] を使用して、300ms の無音区間を境界として取得された発話の回数である。NNS、NS 及びエージェントの発話数を計算する。

2) 話者交替数

会話において、話者間で発話権の移り変わりの回数である NNS、NS 及びエージェント 3 者間の話者交替数 [37] を集計する。また、発話権の移動方向を詳しく調べるため、話者交替数は NS から NNS へ交替、NS からエージェントへ交替、NNS から NS へ交替、NNS からエージェントへ交替、エージェントから NNS へ交替、エージェントから NS へ交替のように分類して、数える。

3) 発話区間

各話者の発話区間は7分の会話において自分が発話する時間の長さである。話者別に発話区間を計算する。

4) 発話・沈黙区間の割合

各話者の沈黙区間は自分の音声がない無音の区間である。話者ごとにそれぞれの発話区間の長さや沈黙区間の長さを測り、各話者の発話・沈黙区間の割合を調べる。

4.2.7 実験手順

実験手順は次の図 18 で示す。

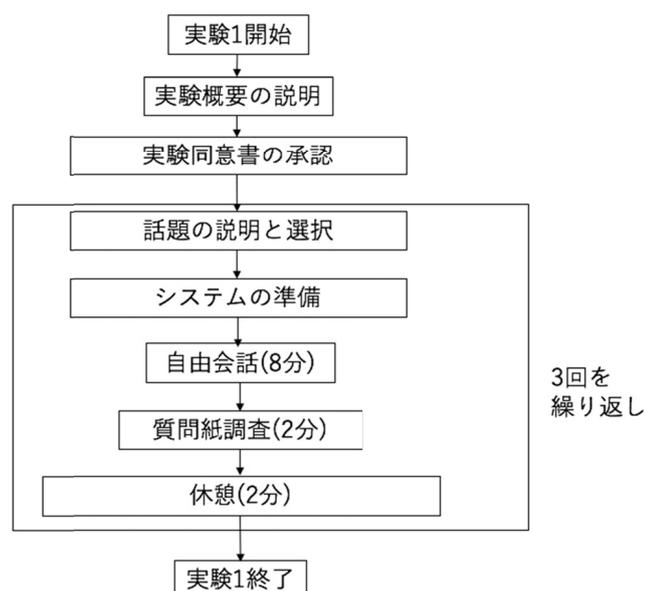


図 18. 実験手順

実験開始前はシステムの設定を完了する。エージェントを含む条件は最初に行われる場合、エージェントを待機画面に変える。

1) 実験開始と実験説明

被検者は入室して椅子に座る。被験者に同意書一部ずつを渡す。実験者は実験の概要、同意書の内容及び実験の手続きについて説明する。説明が終わったら、各被験者には、二枚の同意書に署名してもらい、同意書一枚を回収する。そして、デジタルカメラの撮影を始める。また、マイクをオン”にして、被験者の口に近い場所に装着する。

2) 話題の選択

本実験ではエージェントが NNS-NS 会話への影響を調べて、3つの条件を用意するので、1組は全て違う話題にして、会話を3回行う。会話は自由会話の形式を取る。3回の会話は円滑に進むため、1回の会話は1つの話題について制限なし自由に会話を行ってもらう。日常

的な話題を考え、話題を準備する。それぞれは授業・ゼミ・研究、クラブ活動、趣味、アルバイト、出身、旅行、海外経験、食べ物、普段の生活・習慣、服装、健康、祝日・お祭り・風俗である。会話開始前、実験者は話題の参考として話題用紙を被験者に呈示する。

3) システムの準備

話題選択が完了した後、1 人の実験者は被験者に会話中の注意事項を説明する。そして、被験者に会話開始の合図を送って、会話を開始する。タイマーもスタートし、時間を計る。しかし、タイマーのカウントダウンを被験者に示さない。システム条件では1 人の実験者は注意事項を話す同時に、もう2 人の実験者はシステムの準備を行い、システムを起動する。

4) 自由会話

8 分間の自由会話を行う。8 分になり、タイマーを閉じて、被験者に会話終了の合図を示す。エージェント条件では会話終了時点でシステムも停止する。

5) 質問紙調査と休憩

会話直後に、実験者はNS、NNS にそれぞれのアンケートを配布し、アンケートに記入してもらう。アンケート記入後、2 分間の休憩を入れる。休憩が終わると、次の条件に入る。

6) 会話終了

3 条件の会話が全部終わって全体ビデオを停止し、実験 1 を終了する。実験 1 はおよそ 45 分かかる。

4.3 結果と検討

6 組のデータを得た。各組は3 条件を合わせてデータ処理を行い、エージェント発話条件 42 分、2520 秒、エージェント非発話条件 42 分、2520 秒、及び非エージェント条件 42 分、2520 秒、合計 2 時間 6 分、7560 秒の処理になった。分析項目別に、処理結果の一覧表はエージェント発話条件が表 5 と表 6 で示し、エージェント非発話条件が表 7 と表 8 示し、非エージェント条件が表 9 で示した。

表 5. 処理結果一覧(エージェント発話条件)

エージェント 発話条件	発話数(回)			発話区間(秒)			発話区間割合(%)			沈黙区間(秒)			沈黙区間割合(%)		
	組	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS
1	1.0	96.0	101.0	2.1	251.2	183.7	0.5%	59.8%	43.7%	417.9	168.8	236.3	99.5%	40.2%	56.3%
2	0.0	73.0	125.0	0.0	101.2	112.3	0.0%	24.1%	26.7%	420.0	318.8	307.7	100.0%	75.9%	73.3%
3	0.0	67.0	81.0	0.0	217.1	187.2	0.0%	51.7%	44.6%	420.0	202.9	232.8	100.0%	48.3%	55.4%
4	2.0	114.0	105.0	4.1	195.2	106.6	1.0%	46.5%	25.4%	415.9	224.8	313.4	99.0%	53.5%	74.6%
5	11.0	67.0	140.0	11.1	67.6	126.8	2.7%	16.1%	30.2%	408.9	352.4	293.2	97.3%	83.9%	69.8%
6	1.0	149.0	103.0	3.02	202.6	114.6	0.7%	48.2%	27.3%	417.0	217.4	305.4	99.3%	51.8%	72.7%
平均	2.5	94.3	109.2	3.4	172.5	138.5	0.8%	41.1%	33.0%	416.6	247.5	281.5	99.2%	58.9%	67.0%
SD	4.2	32.6	20.6	4.1	71.7	36.9				4.1	71.7	36.9			

表 6. 処理結果一覧(エージェント発話条件)

エージェント 発話条件	話者交替数 (回)	話者交替の方向(回)					
		NS→NNS	NS→Agent	NNS→NS	NNS→Agent	Agent→NNS	Agent→NS
組							
1	62.0	30.0	0.0	30.0	1.0	0.0	1.0
2	32.0	16.0	0.0	16.0	0.0	0.0	0.0
3	40.0	20.0	0.0	20.0	0.0	0.0	0.0
4	59.0	28.0	1.0	27.0	1.0	0.0	2.0
5	67.0	26.0	4.0	27.0	3.0	4.0	3.0
6	30.0	14.0	1.0	14.0	0.0	1.0	0.0
平均	48.3	22.3	1.0	22.3	0.8	0.8	1.0
SD	16.3	6.6	1.5	6.6	1.2	1.6	1.3

表 7. 処理結果一覧(エージェント非発話条件)

エージェント 非発話 条件	発話数(回)			発話区間(秒)			発話区間割合(%)			沈黙区間(秒)			沈黙区間割合(%)		
	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS
組															
1		98.0	125.0		149.9	223.9		35.7%	53.3%		270.1	196.1		64.3%	46.7%
2		95.0	134.0		124.7	106.5		29.7%	25.3%		295.3	313.5		70.3%	74.7%
3		76.0	95.0		187.5	184.0		44.6%	43.8%		232.5	236.0		55.4%	56.2%
4		122.0	137.0		174.7	113.8		41.6%	27.1%		245.3	306.2		58.4%	72.9%
5		78.0	137.0		74.9	134.7		17.8%	32.1%		345.1	285.3		82.2%	67.9%
6		104.0	145.0		144.6	167.4		34.4%	39.8%		275.4	252.6		65.6%	60.2%
平均		95.5	128.8		142.7	155.0		34.0%	36.9%		277.3	265.0		66.0%	63.1%
SD		17.1	17.8		40.0	45.2					40.0	45.2			

表 8. 処理結果一覧(エージェント非発話条件)

エージェント 非発話条件	話者交替数 (回)	話者交替の方向(回)					
		NS→NNS	NS→Agent	NNS→NS	NNS→Agent	Agent→NNS	Agent→NS
組							
1	89.0	45.0	0.0	44.0	0.0	0.0	0.0
2	45.0	22.0	0.0	23.0	0.0	0.0	0.0
3	47.0	24.0	0.0	23.0	0.0	0.0	0.0
4	50.0	25.0	0.0	25.0	0.0	0.0	0.0
5	55.0	28.0	0.0	27.0	0.0	0.0	0.0
6	41.0	21.0	0.0	20.0	0.0	0.0	0.0
平均	54.5	27.5	0.0	27.0	0.0	0.0	0.0
SD	17.5	8.9	0.0	8.6	0.0	0.0	0.0

表 9. 処理結果一覧(非エージェント条件)

非エージェント 条件	発話数(回)		発話区間 (秒)		発話区間割合 (%)		沈黙区間 (秒)		沈黙区間割合 (%)		話者交 替数 (回)	話者交替の方向 (回)	
	NS	NNS	NS	NNS	NS	NNS	NS	NNS	NS	NNS		NS→ NNS	NNS→NS
組													
1	112.0	163.0	119.1	205.0	28.4%	48.8%	300.9	215.0	71.6%	51.2%	58.0	29.0	29.0
2	142.0	144.0	137.6	123.4	32.8%	29.4%	282.4	296.6	67.2%	70.6%	26.0	13.0	13.0
3	115.0	115.0	154.9	125.0	36.9%	29.8%	265.1	295.0	63.1%	70.2%	28.0	14.0	14.0
4	152.0	82.0	237.8	56.2	56.6%	13.4%	182.2	363.8	43.4%	86.6%	28.0	14.0	14.0
5	61.0	77.0	120.1	190.0	28.6%	45.2%	299.9	230.0	71.4%	54.8%	42.0	21.0	21.0
6	73.0	95.0	103.5	191.3	24.7%	45.6%	316.5	228.7	75.3%	54.4%	27.0	14.0	13.0
平均	109.2	112.7	145.5	148.5	34.6%	35.4%	274.5	271.5	65.4%	64.6%	34.8	17.5	17.3
SD	36.3	34.8	48.5	57.4			48.5	57.4			12.8	6.3	6.5

4.3.1 発話数

発話数については、エージェント発話条件の場合、6組平均としてNSが94.3回であり、NNSが109.2回となった。特にエージェントが2.5回の発話をした。エージェント非発話条件では、NSが95.5回、NNSが128.8回の発話をした。非エージェント条件でNSが109.2回、NNSが112.7回であった(図19)。エージェント非発話条件ではNSとNNSの発話数を対応のあるt検定で分析した結果、有意差がみられた($t(5)=-5.004$, $p<0.01$)。エージェント発話条件と非エージェント条件ではNSとNNSの発話数は有意差が見られなかった。

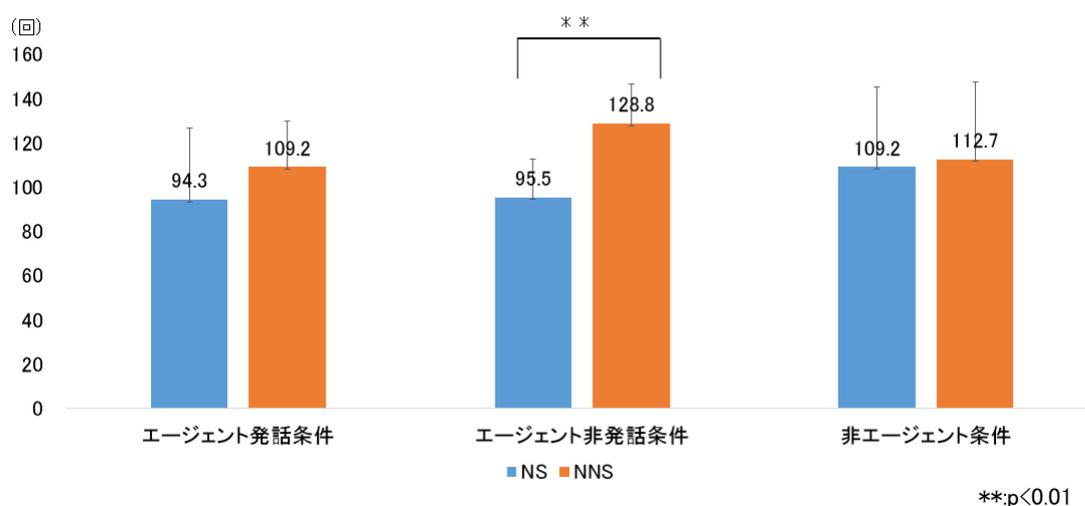


図19. 発話数(7分間)

4.3.2 話者交替数

発話交替の結果は、エージェント発話条件の場合、6組平均で48.3回となった。エージェント非発話条件で54.5回であった。非エージェント条件で34.8回であった(図20)。3条件のfriedman検定を行った結果、有意差は見られた($\chi^2[2]=9.3$, $p<0.01$)。その中、エージェント発話条件と非エージェント条件は有意傾向がみられ($p<0.1$)、エージェント非発話条件と非エージェント条件の間、有意差が見られた($p<0.05$)。

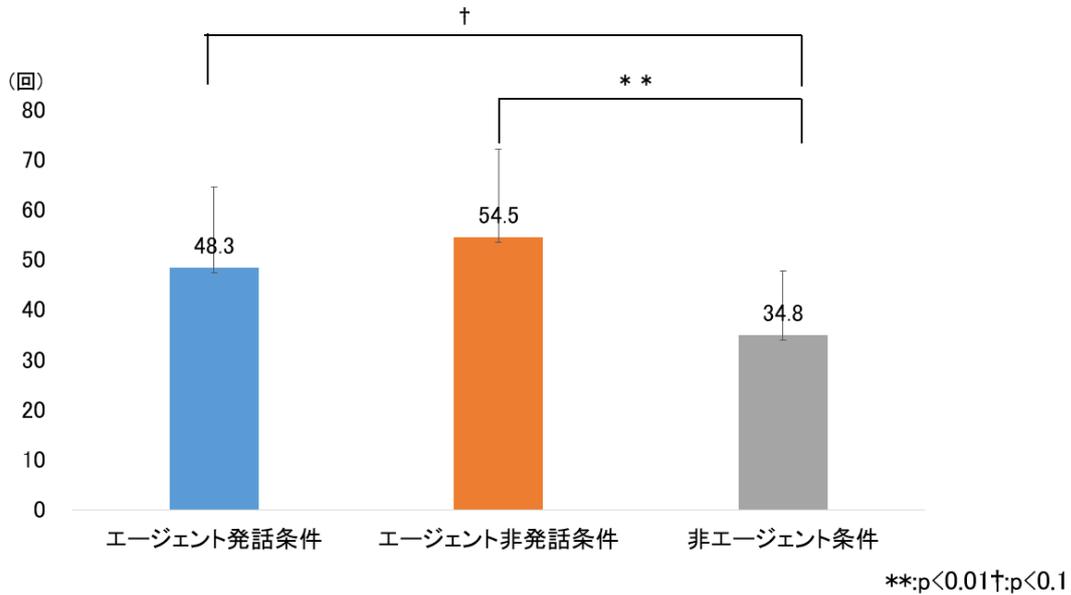


図 20. 話者交替数(7 分間)

また、発話権を取得する回数については、 エージェント発話条件において、 NS から NNS が取った発話権回数が 22.3 回、その反対、 NNS から NS に渡した発話権回数も 22.3 回であった。特に、 NS が Agent に発話権を渡した回数は平均 1.0 回、 NNS が Agent に発話権を渡した回数は平均 0.8 回になった。そして Agent から NNS が得た発話権が 0.8 回、 NS が主導的に Agent から奪った発話権が 1.0 回となった。 エージェント非発話条件において、 エージェントは発話しないため、 Agent から NNS と NS に発話権を渡す回数及び NNS と NS が Agent に発話権を渡す回数は 0 回であった。 NS から NNS が取った発話権が 27.5 回、 NNS から NS に渡した発話権は 27.0 回であった。また、 非エージェント条件において、 NS から NNS が取った発話権回数が 17.5 回、 NNS から NS に渡した発話権回数が 17.3 回であった(図 21) 。 NS から NNS への発話交替数は 3 条件を friedman 検定を行った結果、有意差が見られ ($\chi^2[2]=9.7, p<0.01$)、 エージェント非発話条件は非エージェント条件より有意に多かった ($p<0.01$)。 NNS から NS への発話交替数は friedman 検定の結果、有意差が見られ ($\chi^2[2]=10.2, p<0.01$)、 エージェント発話条件と非エージェント条件の間に有意傾向が ($p<0.1$)、 エージェント非発話条件と非エージェント条件の間に有意差が見られた ($p<0.01$)。

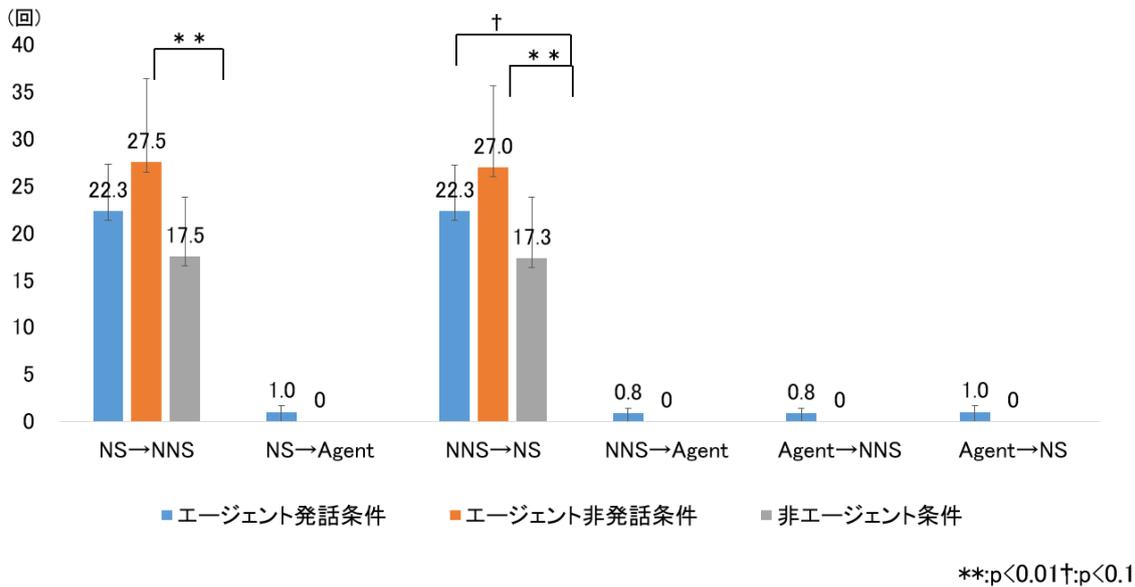


図 21. 話者交替の方向 (7 分間)

4.3.3 発話区間

条件別に話者の発話時間を集計した結果、 エージェント発話条件では、6 組平均で NS が 172.5 秒、 NNS が 138.5 秒となった。 エージェントの発話は 3.4 秒であった。 エージェント非発話条件で NS が 142.7 秒、 NNS が 155.0 秒の発話をした。 非エージェント条件で NS が 145.5 秒、 NNS が 148.5 秒であった (図 22)。 各話者の条件間の発話区間の有意差は見られなかった。 また、 条件別に話者間の有意差も出なかった。

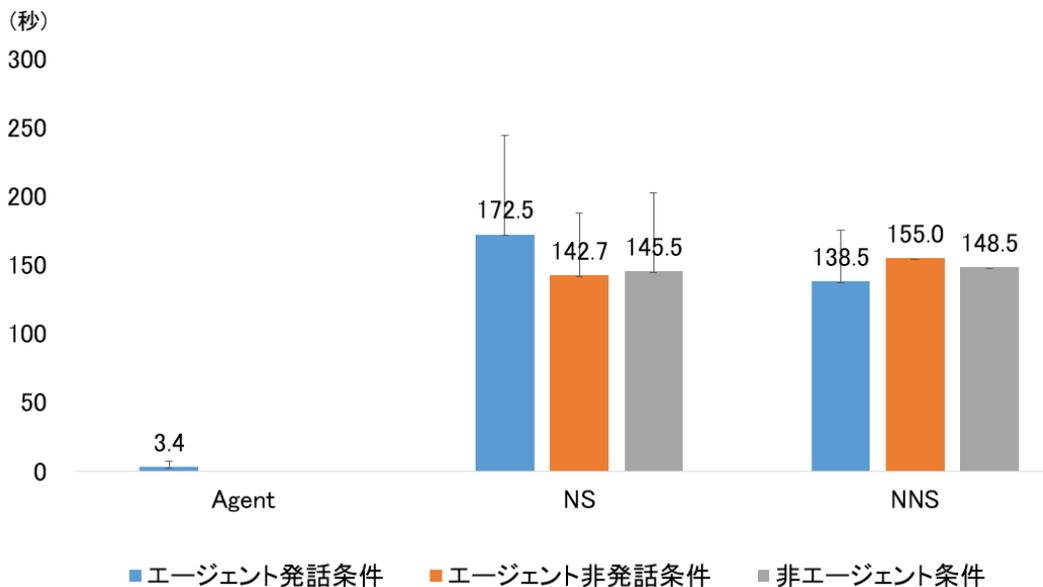


図 22. 発話区間 (7 分間)

4.3.4 発話・沈黙区間の割合

エージェント発話条件の場合，NS の沈黙区間の割合が減少する傾向が見られる．しかし，NNS の沈黙区間の割合は2条件で違いが示されなかった(図 23) ．

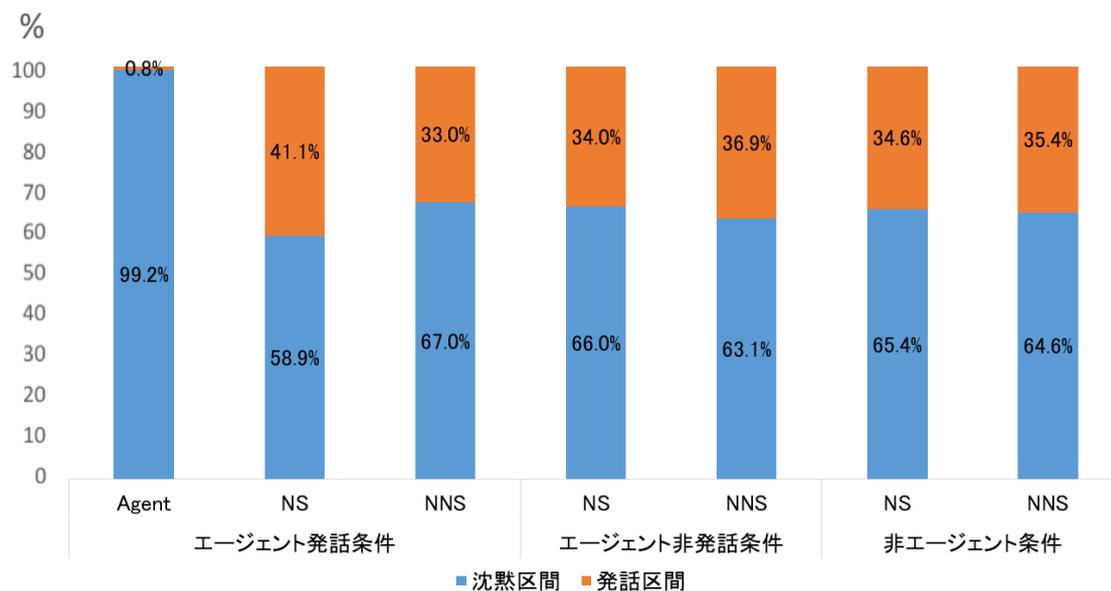


図 23. 発話・沈黙区間の割合

4.3.5 質問紙調査

6 組の質問紙調査の結果は表 10 で示す．NNS と NS それぞれの各項目の得点を 3 条件間で検定した結果，有意差が見られなかった．

表 10. 質問紙調査

番号	質問項目	NNS			NS		
		エージェント ト発話条件	エージェント 非発話条件	非エージェント ト条件	エージェント 発話条件	エージェント 非発話条件	非エージェント 条件
1	自分の言いたい事が話しやすかった。	5.2	5.8	6.0	6.3	6.2	6.3
2	相手と自然にコミュニケーションできた。	5.7	5.8	5.7	6.0	6.3	6.3
3	相手とまた話したいと感じた。	6.0	5.8	5.8	6.2	6.2	6.3
4	言いたいことが遠慮せずに言えた。	5.8	6.0	5.8	6.2	6.0	6.3
5	頻繁に発言することができる。	6.0	5.7	5.8	6.2	6.2	6.3
6	対等な会話ができる。	5.8	5.8	5.7	6.3	6.2	6.5
7	会話が楽しかった。	6.3	5.7	5.7	6.3	6.0	6.3
8	会話に参加している感じがした。	6.2	5.8	5.8	6.3	6.3	6.2
9	会話が盛り上がった。	6.3	6.0	5.7	6.0	6.2	5.8
10	エージェントは会話に参加している感じがした。	4.7	2.8		2.2	1.3	
11	エージェントがいるのは自然だったと思う。	3.7	2.8		3.2	3.7	
12	エージェントは自分(相手)の会話への参加に役立った。	3.7			2.2		
13	エージェントは自然に会話に介入できた。	3.5			2.0		
14	エージェントは自分(相手)の発言を促すことができたと感じた。	3.8			2.7		
15	エージェント介入によって(相手が)多く発言できた。	3.5			2.2		
16	エージェントの発話タイミングが適切だと感じた。	3.5			2.5		
17	相手との会話は不快であった。	1.7	2.2	1.3	1.8	1.7	1.5
18	相手との会話が退屈だと感じた。	2.0	2.0	1.8	1.7	1.8	1.5
19	相手と会話していて疲れた。	2.0	2.2	1.8	2.0	2.2	2.2
20	自分はゆっくりできた。	5.0	5.0	4.8	5.3	5.7	5.2
21	自分は気楽にできた。	5.8	5.7	5.8	5.7	6.0	5.7

4.3.6 検討

1体エージェントを用いて評価実験を行った。発話数の結果について、エージェント非発話条件ではNNSはNSより多く喋ったことが見られた。話者交替の方向の結果から、エージェント非発話条件の場合、NNSとNSはお互いに発話権を渡す回数は非エージェントより多くなった。エージェントの存在はNNSとNS間の話者交替を促すかもしれない。

主観評価については有意差が見られなかったが、NNSの場合は「頻繁に発言することができる」、「会話が楽しかった」、「会話に参加している感じがした」、「会話が盛り上がった」という項目について、エージェント発話条件でより高く評価されたように見える。

「エージェントは会話に参加している感じがした」、「エージェントがいるのは自然だったと思う」という2つの項目について、エージェント発話ありの場面はより会話に参加して、自然の雰囲気をもNNSに伝わったかもしれない。NSの場合はコミュニケーションと会話満足度の面では3条件間で特に大きな違いが示されない。

エージェント非発話条件においてNNSはNSより発話数および話者交替数が多かったため、アンケート結果から会話への介入タイミングや沈黙認識に改善の余地があると思われるが、第二言語会話に役立つことが期待できる。

4.4 2体エージェント

2体のエージェントを会話に介入させることによって、NNSが更により多く発話権を取るかどうかを評価する。

4.5 実験方法

4.5.1 実験環境

実験2は4.2.1節で示しているように、実験1と同じような研究室で行った。実験の環境騒音は40dbである。また、本実験では被験者間の対話距離も図24と図25のように設定する。2体エージェントが対面向きになる場合、NNS、NS及び2体エージェントは隣接している2者が等間隔にし、4者が正方形を囲む。カメラ2台が設置される。1台のカメラでNNSとagent1を撮影し、もう1台はNSとagent2を撮影する(図24)。2体エージェントが隣接になる場合、図で示している。2台のカメラの中、1台はagent1とagent2を撮影し、もう1台はNNSとNSを撮影する(図25)。

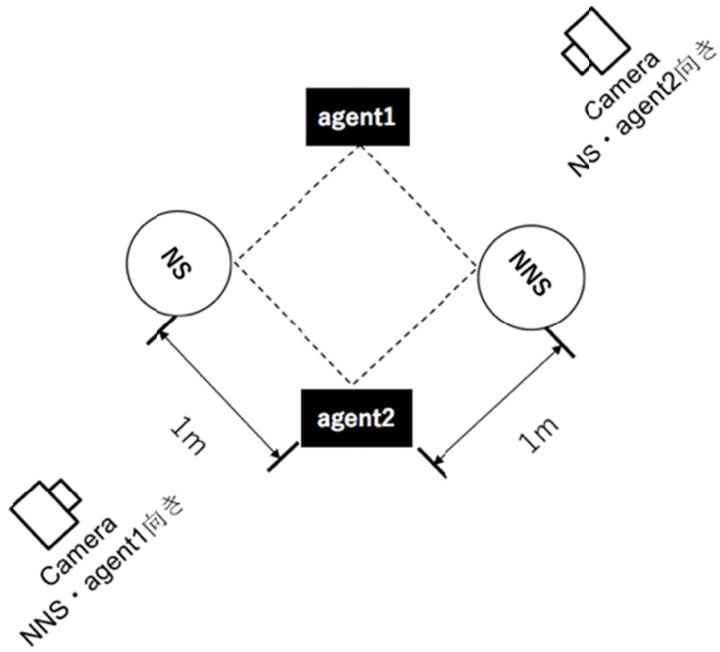


図 24. 実験配置 (2 体エージェント対面向き条件)

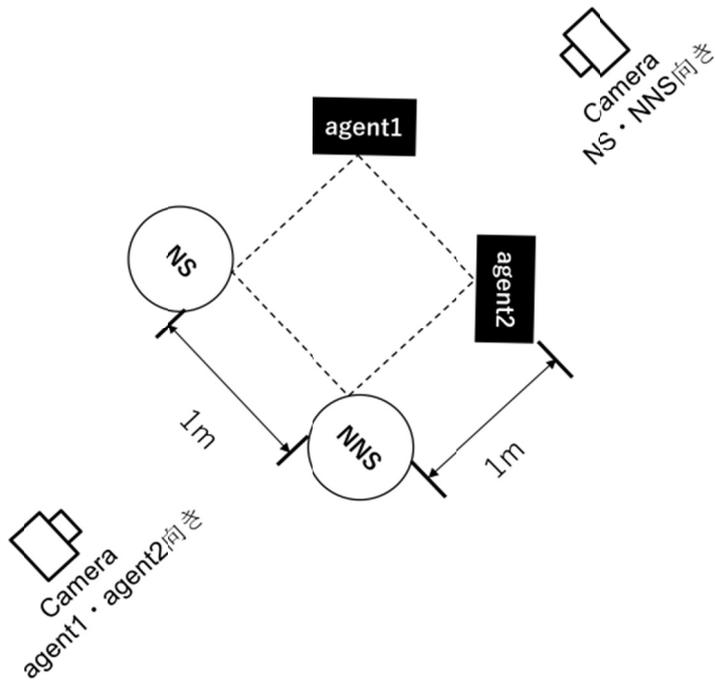


図 25. 実験配置 (2 体エージェント隣接条件)

また、2体エージェント対面向き条件と2体エージェント隣接条件の会話の様子はそれぞれ図26と図27に示される。

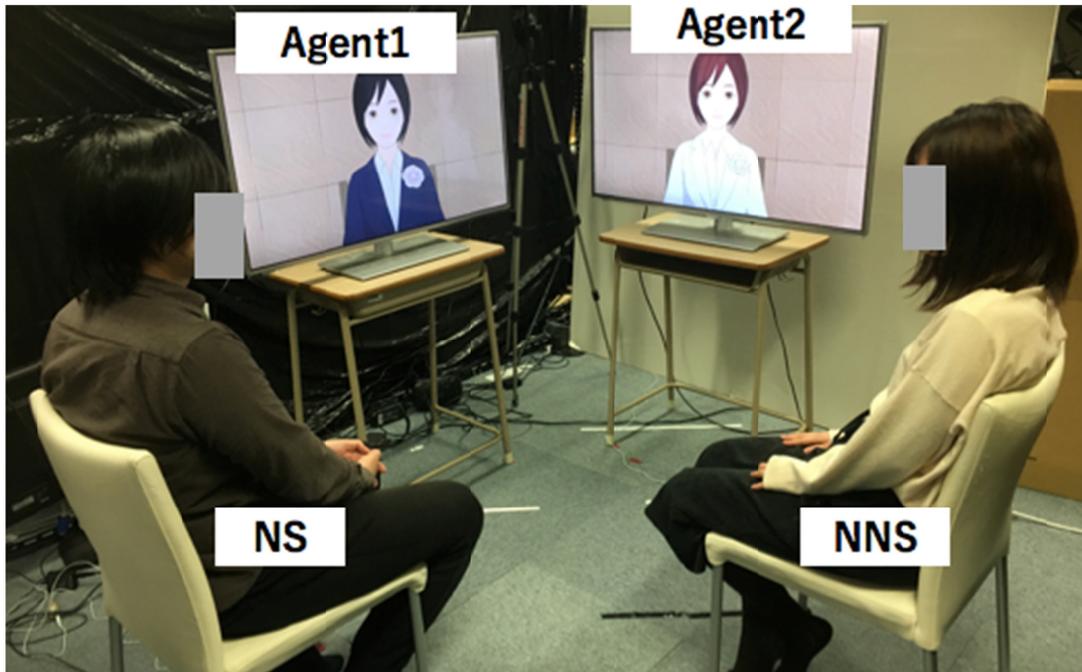


図 26. 実験の様子(2体エージェント対面向き条件)

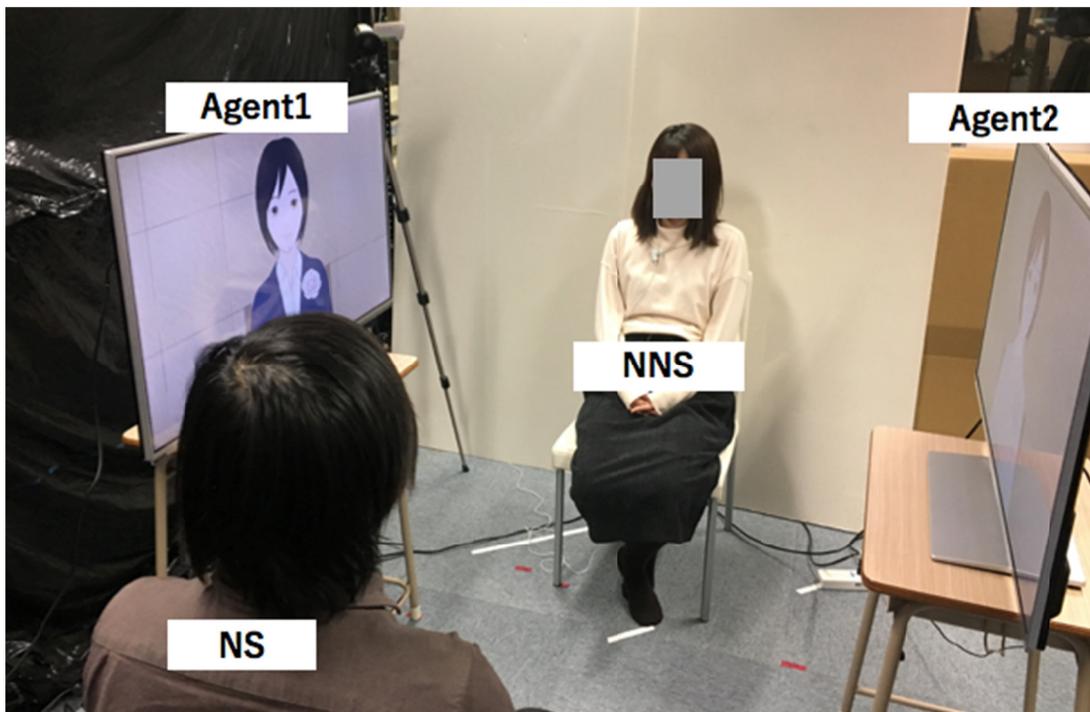


図 27. 実験の様子(2体エージェント隣接条件)

4.5.2 実験装置

4.2.2 で述べた実験装置と同じである。ただし、42 インチのモニターを 2 台使う。

4.5.3 参加者

4.2.3 で述べた参加者である。

4.5.4 実験条件

実験 2 では 1 要因 2 水準の被検者内計画を立てる。

実験の条件は次の 2 条件である。

① 2 体エージェント対面向けの条件

2 体のエージェントは対面向きに位置され、NNS-NS の会話に参加し、発話する。

② 2 体エージェント隣接の条件

2 体のエージェントは隣接しながら、NNS-NS の会話中で発話する。

順序効果を相殺するため、参加者と 2 条件を組み合わせる実験の実施順序は以下のように示す(表 11)。

表 11.実験参加順番

参加者(組)	実験条件	
1	①	②
2	②	①
3	①	②
4	②	①
5	①	②
6	②	①
7	①	②
8	②	①
9	①	②
10	②	①
11	①	②
12	②	①
13	①	②
14	②	①
15	①	②
16	②	①
17	①	②
18	②	①

4.5.5 取得するデータ

4.2.5 で述べた内容と同じようにデータ収集を行う。質問紙について、4.2.5 で述べたものと同じの質問紙を用いる。

4.5.6 データの分析方法

4.2.6 で述べた同じ分析方法で分析項目を用い、データを分析する。

4.5.7 実験手順

実験手順は次の図 28 で示す。

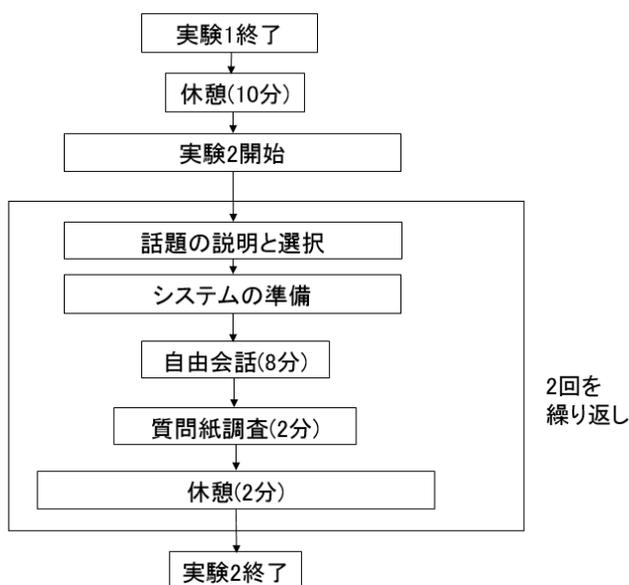


図 28. 実験手順

実験 2 は実験 1 の後に実施する。実験開始前はシステムの設定を完了する。エージェントを待機画面に変える。

1) 実験開始

被検者は椅子に座る。デジタルカメラの撮影を始める。また、マイクを被験者に装着してもらう。

2) 話題の選択

本実験では 4.2.7 で述べた話題用紙と同一のものを使用する。ただし、本実験は実験 1 の後に行い、同じ組の被験者には実験 1 と違う話題を選択してもらう。

3) システムの準備

4.2.7 の 3) で述べた手続きに従う。

4) 自由会話

4.2.7 の 4) で述べた手続きに従う。

5) 質問紙調査と休憩

質問紙の記入は4.2.7の5)と同じ手続きとする。また、2分休憩に入ってから、2体エージェントの位置変換が必要である場合、実験者は休憩の間にエージェントと被験者の位置調整を行う。

6) 会話終了

2条件の会話が全部終わって全体ビデオを停止し、実験2を終了する。

4.6 結果と検討

6組のデータを取得し、データ処理を行った。2体エージェント対面向き条件42分、2520秒、2体エージェント隣接条件42分、2520秒、合計1時間24分、5040秒の処理になった。処理結果の一覧表は2体エージェント対面向き条件が表12と表13で示し、2体エージェント隣接条件が表14と表15示した。

表 12. 処理結果一覧(2体エージェント対面向き条件)

2体エージェント対面向き条件	発話数(回)			発話区間(秒)			発話区間割合(%)			沈黙区間(秒)			沈黙区間割合(%)		
	組	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS
1	0.0	95.0	118.0	0.0	124.5	215.9	0.0%	29.6%	51.4%	420.0	295.5	204.1	100.0%	70.4%	48.6%
2	1	133.0	167.0	2.5	132.2	140.0	0.6%	31.5%	33.3%	417.5	287.8	280.0	99.4%	68.5%	66.7%
3	0	101.0	119.0	0.0	163.2	157.9	0.0%	38.8%	37.6%	420.0	256.8	262.1	100.0%	61.2%	62.4%
4	2	166.0	81.0	1.7	231.8	60.9	0.4%	55.2%	14.5%	418.3	188.2	359.1	99.6%	44.8%	85.5%
5	8	59.0	89.0	6.2	107.0	215.1	1.5%	25.5%	51.2%	413.8	313.0	204.9	98.5%	74.5%	48.8%
6	1	82.0	94.0	2.1	161.0	161.9	0.5%	38.3%	38.6%	417.9	259.0	258.1	99.5%	61.7%	61.4%
平均	2.0	106.0	111.3	2.1	153.3	158.6	0.5%	36.5%	37.8%	417.9	266.7	261.4	99.5%	63.5%	62.2%
SD	3.0	38.1	31.4	2.3	44.1	57.2				2.3	44.1	57.2			

表 13. 処理結果一覧(2体エージェント対面向き条件)

2体エージェント対面向き条件	話者交替数(回)	話者交替の方向(回)											
		NS→ NNS	NS→ Agent1	NS→ Agent2	NNS →NS	NNS→ Agent1	NNS→ Agent2	Agent1 →NNS	Agent1 →NS	Agent1→ Agent2	Agent2 →NNS	Agent2 →NS	Agent2→ Agent1
組													
1	66.0	33.0	0.0	0.0	33.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	29.0	13.0	0.0	0.0	14.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0
3	40.0	20.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	23.0	11.0	0.0	0.0	10.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
5	45.0	19.0	0.0	0.0	18.0	2.0	2.0	0.0	2.0	0.0	2.0	0.0	0.0
6	40.0	19.0	1.0	0.0	19.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
平均	40.5	19.2	0.2	0.0	19.0	0.5	0.5	0.2	0.5	0.0	0.5	0.0	0.0
SD	14.9	7.7	0.4	0.0	7.8	0.8	0.8	0.4	0.8	0.0	0.8	0.0	0.0

表 14. 処理結果一覧(2体エージェント隣接条件)

2体エージェント隣接条件	発話数(回)			発話区間(秒)			発話区間割合(%)			沈黙区間(秒)			沈黙区間割合(%)		
	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS	Agent	NS	NNS
組															
1	0.0	125.0	163.0	0.0	146.5	204.7	0.0%	34.9%	48.7%	420.0	273.5	215.3	100.0%	65.1%	51.3%
2	0.0	58.0	84.0	0.0	148.1	158.1	0.0%	35.3%	37.6%	420.0	271.9	261.9	100.0%	64.7%	62.4%
3	4.0	123.0	109.0	3.8	194.8	127.6	0.9%	46.4%	30.4%	416.2	225.2	292.4	99.1%	53.6%	69.6%
4	0	168.0	70.0	0.0	248.4	47.3	0.0%	59.2%	11.3%	420.0	171.6	372.8	100.0%	40.8%	88.8%
5	9.0	80.0	149.0	9.8	99.0	163.5	2.3%	23.6%	38.9%	410.2	321.0	256.5	97.7%	76.4%	61.1%
6	4.0	123.0	113.0	2.6	153.0	133.3	0.6%	36.4%	31.7%	417.4	267.0	286.7	99.4%	63.6%	68.3%
平均	2.8	112.8	114.7	2.7	165.0	139.1	0.6%	39.3%	33.1%	417.3	255.0	280.9	99.4%	60.7%	66.9%
SD	3.6	38.7	36.0	3.8	51.0	52.6				3.8	51.0	52.6			

表 15. 処理結果一覧(2体エージェント隣接条件)

2体エージェント隣接条件	話者交替数(回)	話者交替の方向(回)											
		NS→ NNS	NS→ Agent1	NS→ Agent2	NNS →NS	NNS→ Agent1	NNS→ Agent2	Agent1 →NNS	Agent1 →NS	Agent1→ Agent2	Agent2 →NNS	Agent2 →NS	Agent2→ Agent1
組													
1	55.0	28.0	0.0	0.0	27.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	26.0	13.0	0.0	0.0	13.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	31.0	13.0	2.0	0.0	14.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
4	23.0	12.0	0.0	0.0	11.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	66.0	28.0	2.0	1.0	28.0	2.0	0.0	2.0	2.0	0.0	1.0	0.0	0.0
6	57.0	26.0	0.0	2.0	27.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0
平均	43.0	20.0	0.7	0.5	20.0	0.3	0.0	0.5	0.5	0.0	0.3	0.2	0.0
SD	18.4	8.1	1.0	0.8	8.1	0.8	0.0	0.8	0.8	0.0	0.5	0.4	0.0

4.6.1 発話数

2体エージェント対面向き条件で、6組平均としてNSが106.0回であり、NNSが111.3回となった。エージェントが2.0回の発話をした。2体エージェント隣接条件では、NSが112.8回、NNSが114.7回であった(図29)。条件別にNSとNNSの発話数を対応のあるt検定で分析した結果、有意差が見られなかった。

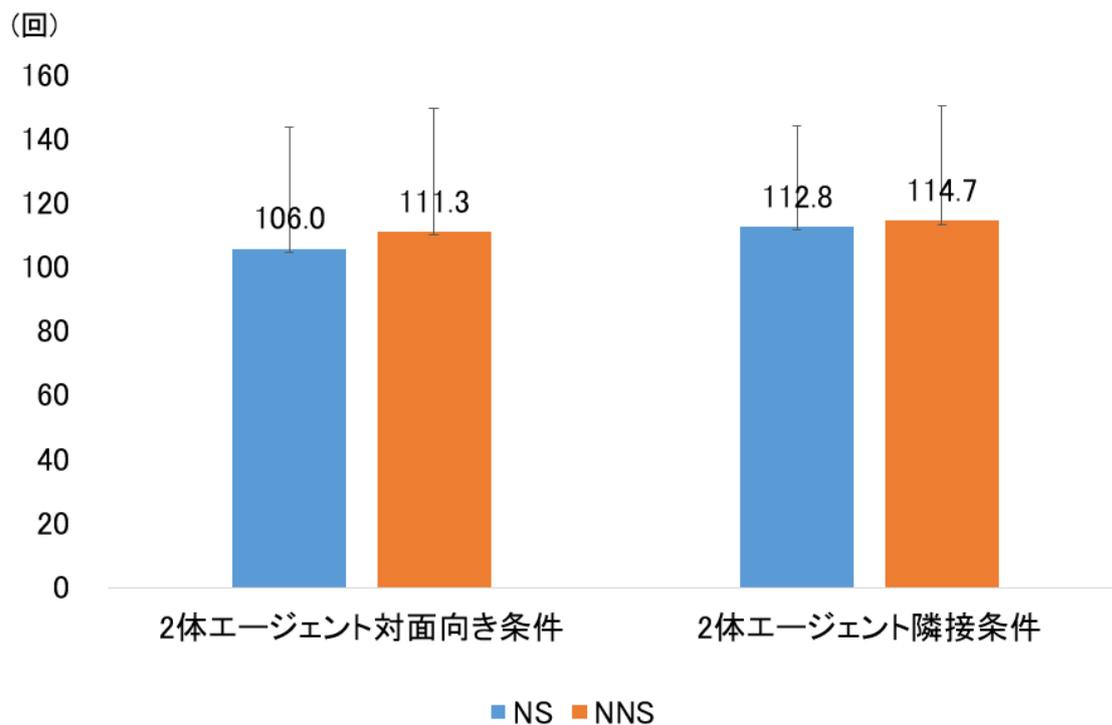


図 29. 発話数(7 分間)

4.6.2 話者交替数

発話交替の結果は、2 体エージェント対面向き条件の場合、40.5 回となった。2 体エージェント隣接条件で 43.0 回であった(図 30)。2 条件は対応のある t 検定で分析した結果、有意差が見られなかった。

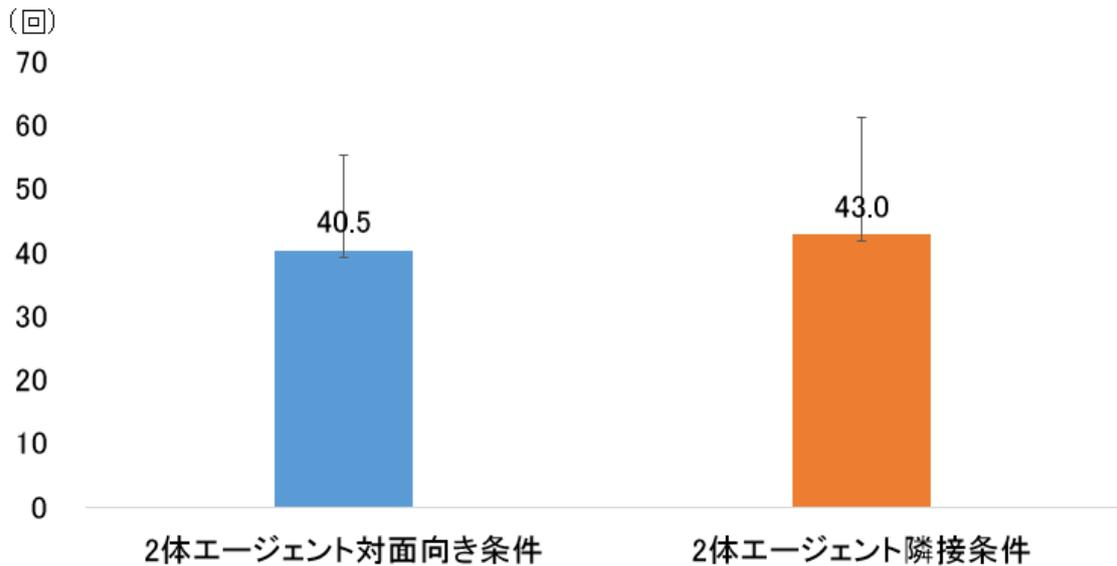


図 30. 話者交替数(7 分間)

また、発話権を取得する回数については、2体エージェント対面向き条件の場合、NS から NNS が取った発話権回数が 19.2 回、NNS から NS に渡した発話権回数が 19.0 回であった。2体エージェントがいるため、NS は Agent1 に発話権を渡した回数が 0.2 回、Agent2 に渡す発話権の回数が 0 回であった。NNS から Agent1 に発話権を渡した回数は 0.5 回、Agent2 に渡す発話権の回数も 0.5 になった。また、Agent1 から NNS が得た発話権が 0.2 回、NS が主導的に Agent1 から奪った発話権が 0.5 回となった。Agent2 から NNS が得た発話権が 0.5 回、NS が主導的に Agent2 から奪った発話権が 0 回であった。Agent1 から Agent2 または Agent2 から Agent1 に発話権を渡す場面はなかった。2体エージェント隣接条件において、NS から NNS が取った発話権回数が 20.0 回、NNS から NS に渡した発話権回数が 20.0 回であった。そして、NS は Agent1 に発話権を渡した回数が 0.7 回、Agent2 に渡す発話権の回数が 0.5 回であった。NNS から Agent1 に発話権を渡した回数は 0.3 回、Agent2 に渡す発話権の回数は 0 回であった。また、Agent1 から NNS が得た発話権が 0.5 回、NS が主導的に Agent1 から奪った発話権が 0.5 回となった。Agent2 から NNS が得た発話権が 0.3 回、NS が主導的に Agent2 から奪った発話権が 0.2 回になった。Agent1 から Agent2 または Agent2 から Agent1 に発話権を渡す場面はなかった(図 31)。NNS が NS からもらえた発話権の回数及び NNS から NS に渡す発話権の回数は wilcoxon の符号付き順位検定で分析した結果、有意差は見られた。

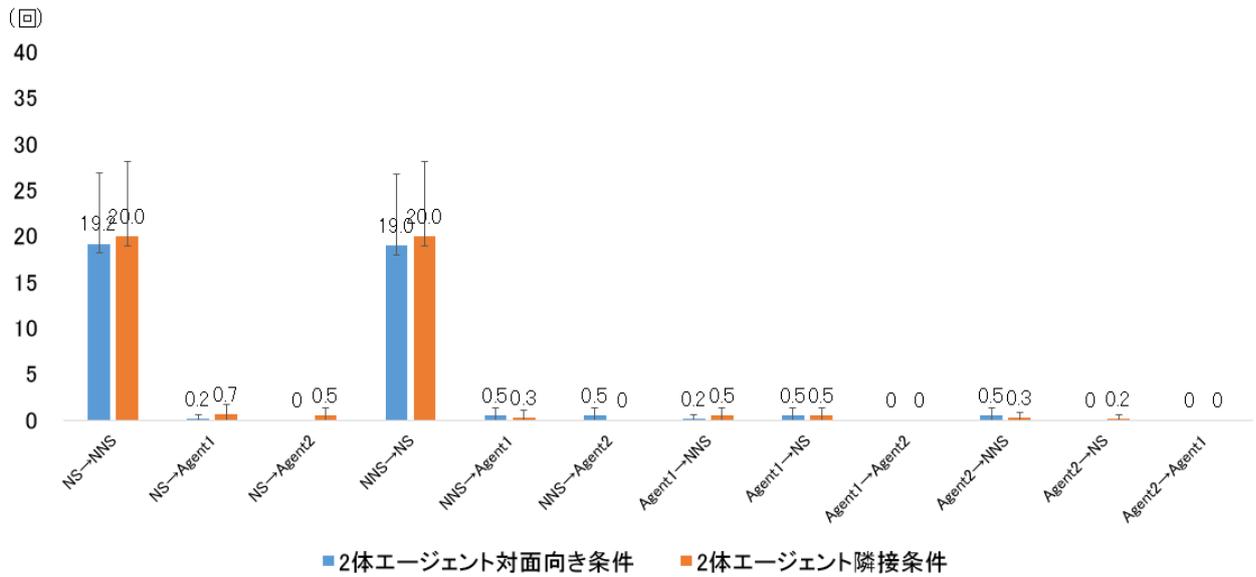


図 31. 話者交替の方向 (7 分間)

4. 6. 3 発話区間

発話区間について 2 体エージェント対面向き条件の場合、6 組平均で NS が 153.3 秒、NNS が 158.6 秒となった。エージェントの発話は 2.1 秒であった。2 体エージェント隣接条件で NS が 165.0 秒、NNS が 139.1 秒であった(図 32)。各話者の条件間の発話区間の有意差は見られなかった。また、条件別に話者間の有意差も出なかった。

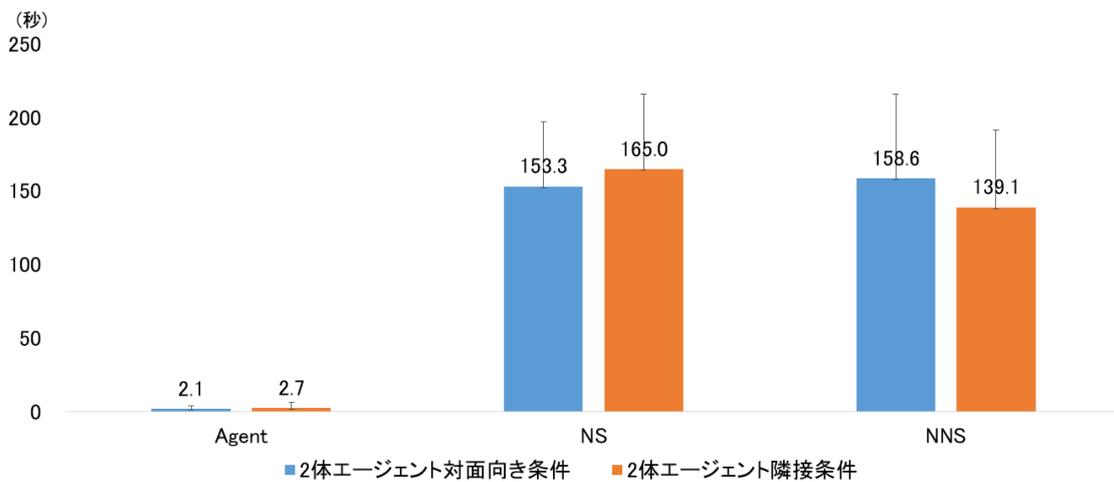


図 32. 発話区間 (7 分間)

4. 6. 4 発話・沈黙区間の割合

図 34 は 2 条件で各話者の発話と沈黙区間の割合が示される。2 体エージェント対面向き条件と 2 体エージェント隣接条件の間で、NS の発話・沈黙区間の割合が特に大きな違いが見られなかった。NNS の場合、対面向き条件のほうは隣接条件より発話区間の割合が増える傾向が見られた(図 33)。

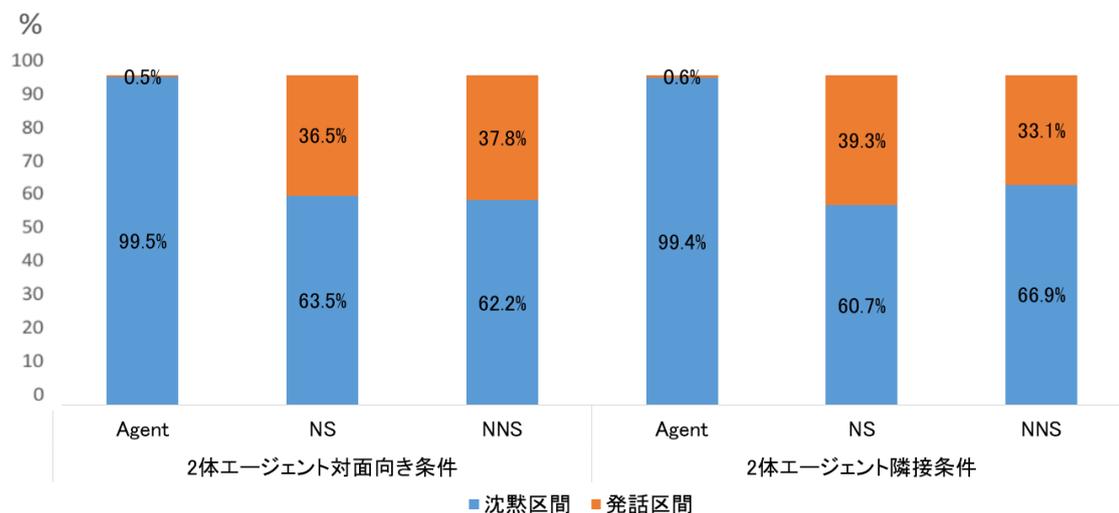


図 33. 発話・沈黙区間の割合

4.6.5 質問紙調査

6 組の質問紙調査の結果は表 16 で示す。NNS と NS それぞれの各項目の得点を条件間で検定した結果、有意差が見られなかった。

表 16. 質問紙調査

番号	質問項目	NNS		NS	
		2体エージェント対面向き条件	2体エージェント隣接条件	2体エージェント対面向き条件	2体エージェント隣接条件
1	自分の言いたい事が話しやすかった。	6.2	6.0	6.0	6.2
2	相手と自然にコミュニケーションできた。	6.3	5.8	6.2	6.3
3	相手とまた話したいと感じた。	6.2	6.0	6.3	6.5
4	言いたいことが遠慮せずに言えた。	6.2	5.7	6.2	6.5
5	頻繁に発言することができる。	6.2	5.8	6.0	6.3
6	対等な会話ができる。	6.0	5.7	6.3	6.2
7	会話が楽しかった。	6.2	6.0	6.2	6.5
8	会話に参加している感じがした。	6.3	5.8	6.3	6.5
9	会話が盛り上がった。	6.2	6.0	6.5	6.2
10	エージェントは会話に参加している感じがした。	3.8	3.5	2.7	1.7
11	エージェントがいるのは自然だったと思う。	3.3	2.8	3.0	2.7
12	エージェントは自分(相手)の会話への参加に役立った。	3.2	2.5	2.6	1.8
13	エージェントは自然に会話に介入できた。	3.3	2.8	1.8	1.7
14	エージェントは自分(相手)の発言を促すことができたと感じた。	3.7	3.3	2.2	2.3
15	エージェント介入によって(相手が)多く発言できた。	3.3	3.3	2.7	1.8
16	エージェントの発話タイミングが適切だと感じた。	2.8	3.2	2.7	1.7
17	相手との会話は不快であった。	1.3	1.8	1.5	1.3
18	相手との会話が退屈だと感じた。	1.8	2.0	1.7	1.5
19	相手と会話していて疲れた。	1.8	2.2	1.5	1.5
20	自分はゆっくりできた。	5.0	5.2	5.7	6.0
21	自分は気楽にできた。	6.3	6.0	6.2	6.3

4.6.6 検討

2体エージェントの評価実験を行った。発話数の結果について、エージェント対面向き条件とエージェント隣接条件の間にNNSとNSの発話数は違いが見られなかったため、2体エージェントの位置は話者の発話数に影響を与えないと考えられる。話者交替について2体エージェントの場合、条件間で違いがなかった。また、話者交替の方向の結果として、NSからNNSへの話者交替回数およびNNSからNSへの話者交替回数は条件間で違いが示されなかった。NNSとNS間の話者交替は2体エージェントの位置関係に影響されなかったと考えられる。主観評価の結果についても、有意差が見られなかった。

今後の課題として、1体エージェントの効果との比較が必要と考えられる。

第5章 結論

本研究では第二言語会話における NNS の発話機会の少なさに対し、NNS に発話権を渡すことができるようなエージェントを提案する。提案システムについては NNS-NS の会話中でエージェントが適切なタイミングで NNS に発話権を渡すことができる。システムの有用性を検証するため、評価実験を設計する。6 組の初期実験の結果を示した。エージェントがいる場合、NNS が多く発言でき、NNS-NS 間の話者交替が増加したことが見られた。また、エージェントの数を増え、2 体を会話に介入させる場合も提案した。今後、1 体エージェントの効果との比較が必要と考えられる。

謝辞

本修士論文は、筆者が筑波大学大学院図書館情報メディア研究科図書館情報メディア専攻博士前期課程にて、井上研究室で行った研究をもとめたものである。本研究に関して終始ご指導ご鞭撻を頂きました指導教員井上智雄教授に心より深く感謝致します。また、本論文をご精読頂き有用なコメントを頂きました副指導教員森田ひろみ准教授に深謝致します。

更に、最後まで一緒に頑張ってきて多くのご指摘を下さいました同研究室の同期・後輩の皆様及び実験の際に、被験者を快く引き受けてくださり、実験に協力していただいた方々に心より感謝しております。ありがとうございました。

参考文献

- [1] Zuengler, Jane. "Encouraging learners' conversational participation: The effect of content knowledge." *Language Learning* 43(3) pp.403-432 (1993).
- [2] Van Lier, Leo, Naoko Matsuo. "Varieties of conversational experience: Looking for learning opportunities." *Applied language learning* 11(2) pp.265-288 (2000).
- [3] Zhu, Wei. "Interaction and feedback in mixed peer response groups." *Journal of second language writing* 10(4) pp.251-276 (2001).
- [4] 一二三朋子. "非母語話者との会話における母語話者の言語面と意識面との特徴及び両者の関連." *教育心理学研究* 47(4) pp.490-500 (1999).
- [5] Gao, G., Yamashita, N., Hautasaari, A. M., & Fussell, S. R. Improving multilingual collaboration by displaying how non-native speakers use automated transcripts and bilingual dictionaries. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems -CHI' 15*, Seoul, Republic of Korea, April18-23, 2015, pp. 3463-3472 ACM.
- [6] Hautasaari, Ari, Naomi Yamashita. "Catching up in audio conferences: highlighting keywords in ASR transcripts for non-native speakers." *Proceedings of the 5th ACM international conference on Collaboration across boundaries: culture, distance & technology-CABS' 14*, Kyoto, Japan, August 20 - 22, 2014, pp. 107-110 ACM.
- [7] Hanawa, Hiromi, Xiaoyu Song, Tomoo Inoue. "Key-Typing on Teleconference: Collaborative Effort on Cross-Cultural Discussion." *International Conference on Collaboration Technologies*. vol647, pp. 74-88, 2016, Springer Singapore.
- [8] Ye, Jing, Tomoo Inoue. "A Speech Speed Awareness System for Non-Native Speakers." *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion-CSCW' 16 Companion*, San Francisco, California, USA, February 26 - March 02, 2016, pp. 49-52, ACM.
- [9] 藤本学. "会話者のコミュニケーション参与スタイルを指し示す COMPASS." *社会心理学研究* 23(3) pp.290-297 (2008).
- [10] 藤本学, 大坊郁夫. "小集団による会話の展開に及ぼす会話者の発話行動傾向の影響." *実験社会心理学研究* 47(1) pp.51-60 (2007).
- [11] Fujimoto, M., Murayama, A, & Daibo, I. (August 2004) What ENSURES a small-group discussion efficient? (1): The influence of group structure on

- various aspects of the small-group discussion. XXV International Congress of Psychology. Beijing, CHINA.
- [12] 藤本学, 大坊郁夫. "小集団会話における話者の発言傾向を規定する3要素 (特集) 実験による言語行動の研究)." 社会言語科学 9(1) pp. 48-58 (2006).
- [13] 杉山貴昭, 船越孝太郎, 中野幹生, 駒谷和範. "多人数対話におけるロボットの応答義務の推定." 人工知能学会全国大会論文集 29, 1-4, 2015.
- [14] 八城美里, 林佑樹, 中野有紀子. "会話参加者の優位性を考慮した会話エージェントによる多人数会話への介入." 人工知能学会全国大会論文集 27 1-4, 2013.
- [15] 黄宏軒, 乙木翔地, 堀田怜, 川越恭二. "多人数会話において積極的に情報提示ができるガイドエージェントの実現に向けての介入場面の検討." 人工知能学会論文誌 31(1), (2016), DSF-514, pp. 1-13.
- [16] 山内祐輝, 戸田智基, 中村哲. "対話システムにおける用語間の関係性を用いた話題誘導応答文生成." 人工知能学会論文誌 29(1), (2014), pp. 80-89.
- [17] 轡田真治, 井上智雄. "多人数会話における動的な会話支援システムの開発." 研究報告グループウェアとネットワークサービス 2009-GN-72(6), 1-6. (2009).
- [18] 高瀬裕, 吉野堯, 中野有紀子. "多人数会話における調整・介入機能を有する対話ロボット." 情報処理学会論文誌 58(5) pp. 967-980 (2017).
- [19] 秋葉巖, 松山洋一, 小林哲則. "多人数会話ファシリテーションロボットの主導権奪取手続き." 研究報告音声言語情報処理 (SLP) 2013-SLP-97(10), 1-8, 2013-07-18.
- [20] Benne, Kenneth D., Paul Sheats. "Functional roles of group members." Journal of social issues 4.2 (1948): 41-49.
- [21] 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-インタラクション行動から興味対象を抽出する" 認知科学, 11(3), pp. 214-227, (2004).
- [22] <http://www.mmdagent.jp/>
- [23] Lee, A., Oura, K., Tokuda, K. MMDAgent-A fully open-source toolkit for voice interaction systems. 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013), pp. 8382-8385. IEEE. (2013, May).
- [24] <https://people.csail.mit.edu/hubert/pyaudio/docs/>
- [25] 切替一郎, 戸塚元吉, 岡本途也, 河村正三, 佐藤恒正, 設楽哲也, 般坂宗太郎. 聴覚検査法. 医学書院. (1974).
- [26] 塩原拓人, 井上智雄. "共食エージェントがユーザの食事に及ぼす影響." 研究報告デジタルコンテンツクリエーション (DCC) 2013-DCC-4(12), 1-8, (2013).
- [27] サックス, シェグロフ, ジェファンソン, 西阪仰 訳 "会話分析基本論集 順番

- 交替と修復の組織。”世界思想社（2010）。
- [28] Nakano, Yukiko, Yuki Fukuhara. “Estimating conversational dominance in multiparty interaction.” Proceedings of the 14th ACM international conference on Multimodal interaction-ICMI’ 12, Santa Monica, California, USA, October 22-26, 2012, pp.77-84. ACM.
- [29] <http://open-jtalk.sp.nitech.ac.jp/>
- [30] Rubin, J. (1994). A review of second language listening comprehension research. The modern language journal, 78(2), 199-221.
- [31] 岡本健吾, 吉野孝. “J-005 可視化した会話中のキーワードを用いた対面型異文化間コミュニケーション支援システムの開発 (ヒューマンコミュニケーション & インタラクション, 一般論文).” 情報科学技術フォーラム講演論文集 8.3 (2009): 393-396.
- [32] RuiLiu Tomoo Inoue Application of An Anthropomorphic Dining Agent to Idea Generation In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication-UBICOMP ’14 ADJUNCT, SEATTLE, WA, USA, SEPTEMBER 13 - 17, 2014, pp607-612. ACM.
- [33] Tomoo Inoue Naturalistic Control of Conversation by Meal: Induction of Attentive Listening Attitude through Uneven Meal Distribution in Co-dining. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication- UBIComp ’14 ADJUNCT, SEATTLE, WA, USA, SEPTEMBER 13 - 17, 2014, pp601-606. ACM.
- [34] 高橋徹, 濱崎雅弘, 武田英明. (2001). Avatar-like エージェントを用いた Web コミュニティ支援システム. 人工知能学会全国大会論文集, (0), 80-80.
- [35] <https://tla.mpi.nl/tools/tla-tools/elan/>
- [36] 大須賀智子, 堀内靖雄, 西田昌史, 市川薫”音声対話での話者交替/継続の予測における韻律情報の有効性.” 人工知能学会論文誌 21(1), (2006), pp.1-8.
- [37] 榎本美香. “会話・対話・談話研究のための分析単位: ターン構成単位 (TCU) (< 連載チュートリアル> 多人数インタラクションの分析手法 [第 4 回]).” 人工知能学会誌 23(2), (2008), pp.265-270.