

エージェントによる第二言語会話支援の
研究

筑波大学

図書館情報メディア研究科

2019年3月

GUO ZIXUAN

目次

第1章	序論	4
1.1	研究背景	4
1.2	研究目的	4
1.3	構成	4
第2章	関連研究	5
2.1	対面会話における基本規則	5
2.1.1	参与役割(Participant Role)	5
2.1.2	話者交替(Turn-taking)	6
2.1.3	修復(Repair)	7
2.2	擬人化会話エージェント	9
2.2.1	単体エージェントによる会話	9
2.2.2	複数体エージェントによる会話	11
2.3	本研究の位置づけ	13
第3章	NNS 支援エージェントの概要と評価実験	14
3.1	NNS 支援エージェントの概要	14
3.2	参加者と条件	16
3.3	会話タスク	17
3.4	環境と手順	17
3.5	データ収集	18
3.5.1	会話行動	18
3.5.2	質問紙調査	18
3.6	データ処理	18
3.6.1	データ量	18
3.6.2	発話区間のラベリング	18
3.6.3	ターン(発話権)の判別とラベリング	19
3.6.4	介入分類のラベリング	20
3.6.5	介入前後の会話内容の分析	22
3.6.6	内容語のラベリング	24
第4章	実験結果と考察	24
4.1	発話頻度	24
4.2	発話量	25
4.3	発話長	26
4.4	あいづち率	27

4.5	発話均衡度.....	28
4.6	話者交替頻度.....	29
4.7	介入の分類.....	30
4.8	介入前後の会話内容の分析.....	31
4.9	介入影響度.....	31
4.10	質問紙調査.....	32
4.11	結論.....	34
第5章	複数体エージェントによる第二言語会話支援の検討.....	35
5.1	提案手法.....	35
5.2	話者交替パターンの設計.....	35
5.2.1	A1：発話権譲渡役.....	35
5.2.2	A2：修復役.....	36
5.3	発話内容の設計.....	36
5.4	プロトタイプと予備実験.....	37
第6章	結論.....	39
	謝辞.....	40
	参考文献.....	41

図表の目次

図 1	対面会話における参与役割.....	5
図 2	話者交替規則.....	6
図 3	会話エージェントを用いた二者会話.....	10
図 4	会話エージェントを用いた多人数会話.....	10
図 5	複数体エージェントを用いた多人数会話.....	12
図 6	本研究の位置づけ.....	14
図 7	NNS 支援エージェントを用いた第二言語会話.....	15
図 8	エージェントの介入タイミング.....	16
図 9	実験の風景.....	16
図 10	実験の配置.....	17
図 11	発話区間ラベリングの例:会話の映像データ (A), 会話音声データの波形 (C), NNS・NS・Agent それぞれの発話区間のラベリング (D) ラベリングした発話区間 のリスト (B).....	19
図 12	ターンラベリングの例:ELAN で作成した発話区間のラベリング (上), 画像 編集ソフトで作成したターンのラベリング (下), NNS・NS・Agent それぞれのタ	

ーン (赤色・緑色・黄色の部分), あいづち (紫色の丸で囲まれている部分) .	20
図 13 発話頻度(N = 18)	25
図 14 発話量(N = 18).....	26
図 15 発話長(N = 18).....	27
図 16 あいづち率 (N = 18).....	28
図 17 発話均衡度 (N = 18).....	29
図 18 話者交替頻度 (N = 18)	29
図 19 介入影響度(DS と IS: N = 22, RF: N = 28).....	32
図 20 質問紙調査(1/2) (N = 18).....	33
図 21 質問紙調査(2/2) (N = 18).....	34
図 22 システム構成	37
図 23 予備実験の様子	38
表 1 エージェントの発話内容	15
表 2 DS の会話例	21
表 3 IS の会話例	21
表 4 RF の会話例	22
表 5 NF の会話例	22
表 6 日本語における他者開始形式 (鈴木[33])	23
表 7 介入直後に他者開始が発生した例	24
表 8 エージェント介入の分類結果	30
表 9 A2 の発話内容.....	36
表 10 提案エージェントによる会話の一例.....	37
表 11 インタビューの質問項目	39

第1章 序論

1.1 研究背景

グローバル化に伴い、母語が異なる人々の間でコミュニケーションを行う機会が増えている。非母語話者 (Non-native Speaker, 以下 NNS) が母語話者 (Native Speaker, 以下 NS) と会話を行う場合、NNS の第二言語運用力は NS に比べて弱い。その言語運用能力の低さにより、NNS と NS が対等に会話することが難しいことがある[1]。

特に、NS-NNS の会話いわゆる第二言語会話の観察結果[2]によると、NNS は、NS と同じ程度の知識を持っている場合であっても、NS に比べて会話への参加度が低いことが分かった。また、NS と NNS のグループディスカッションにおいて、NNS は、同じグループの NS より発言権を取るのが少ないことが実験によって示された[3]。要するに、NNS は発言に困難を抱え、特に発言権の取得に困難を抱えている。

このような状況において、協調的な第二言語会話を実現するために、NNS がより積極的に会話に参加できる仕組みが望まれる。そのため、我々は話者交替規則に基づく NNS 支援エージェント[4][5]を提案している。NNS 支援エージェントは第二言語会話に介入し、NNS に対して意見を伺うような発言をすることで、NNS の発言を促すことを目標としている。初期評価では、エージェントがいる場合に会話のやり取りが頻繁になる傾向が見られたが、エージェントの介入効率や参加者の会話の変化などが十分に検討されていない。

1.2 研究目的

本研究は、NNS の発言における困難さに着目し、会話エージェントによる第二言語会話支援を目的としている。本稿では(a)NNS 支援エージェントの評価実験における、エージェントの介入効率や参加者の会話の変化などについて述べ、エージェントの有効性と課題を考察する。(b)評価実験から得られた知見を踏まえ、複数体エージェントによる第二言語会話支援手法を検討する。

1.3 構成

本稿は、本章を含め 6 つの章で構成される。第 1 章において研究背景と目的を述べ、第 2 章において関連研究と本研究の位置づけを紹介する。第 3 章では NNS 支援エージェントの概要と評価実験について、第 4 章では実験結果と考察を述べる。第 5 章では複数体エージェントによる第二言語会話支援手法の検討とプロトタイプを述べ、第 6 章を結論とする。

第2章 関連研究

近年，擬人化会話エージェント（Life-like Agent），ECA (Embodied Conversational Agent) と呼ばれるソフトウェアエージェントが普及してきた。人間であるユーザと擬人化エージェントの間に，多種多様なインタラクションが生まれる。そういうインタラクションの設計問題は HAI(Human-Agent Interaction) と呼ばれる。ある具体的な課題を解決していくにはどのようなインタラクションを設計すればいいのかについて，様々な研究が進められており，特に対話タスクの達成を目的としたエージェントが活発に議論されている。

人同士の対面会話において，社会的な規則が存在するため，社会性を持った擬人化会話エージェントがその仕組みに従う考慮が必要である。

それゆえに，本章では，まず 2.1 節において，対面会話の基本的な規則を紹介する。次に，それらの規則を着眼点として，2.2 節において擬人化会話エージェントの関連研究を紹介し，2.3 節において本研究の位置づけを述べる。

2.1 対面会話における基本規則

2.1.1 参与役割(Participant Role)

対面会話において，発話をしている人を話し手(Speaker)と呼び，発話を受けている人を聞き手(Addressee)と呼ぶ。参加者が相互に話し手と聞き手という役割を交替しながら会話を行う。先行研究[6]において，このような役割の配分が「参与役割」という言葉で表現されている。

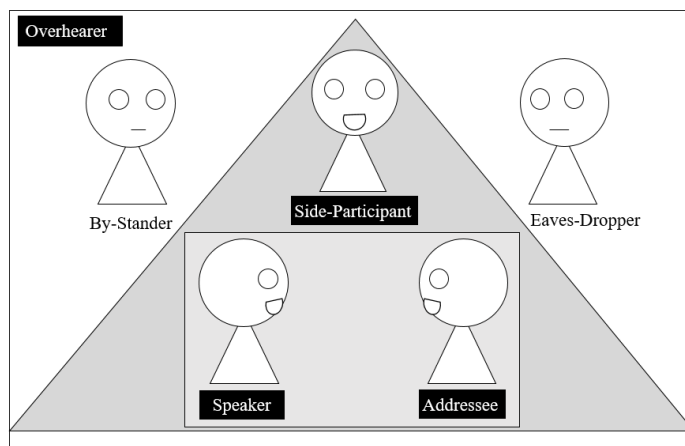


図 1 対面会話における参与役割

図 1 で示すように，話し手(Speaker)に承認されている参加者として，発話が宛てられている人を「聞き手(Addressee)」と定義し，参与はしているが発話が宛てられていない人を

「傍参与者(Side-participant)」と定義する。また、話し手に承認されていない参加者を「傍観者 (Bystander)」と、発話を偶然に聞いてしまった人を「盗み聞きしている者 (Eavesdropper)」と定義する。

社会的存在として認識される擬人化エージェントは、人とインタラクションを取る際に、多様な参加役割を担うことができる。従って、会話の成り行きをコントロールする可能性もある。

2.1.2 話者交替(Turn-taking)

話者交替とは、話し手の移り変わりを指す。Sacks ら[7]は、自然会話の録音データに基づき、話者交替(Turn-taking)の最も単純かつ組織的な形をまとめた (図 2)。

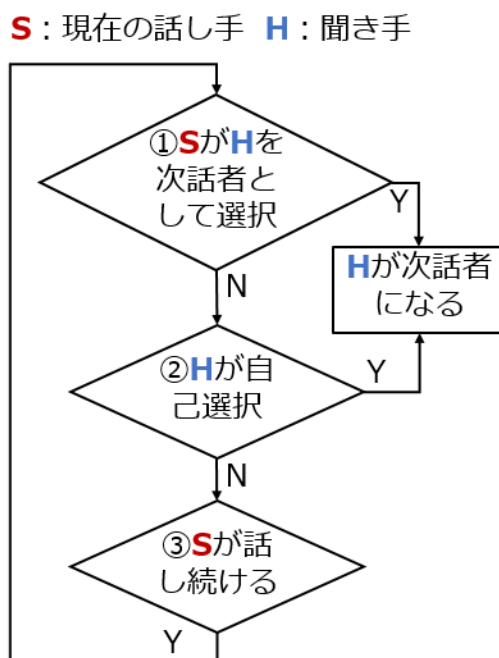


図 2 話者交替規則

話し手が話し始めるから、次の話し手(Next-speaker)と交替するまでの発話のことをターン(Turn)と呼ぶ。話し手が交替するとき、誰が次の話し手になるのかという、その話し手の割り当てに関するルールを話者交替規則と呼び、以下の3種類から構成される。

① 他者選択規則

現在の話し手が次話者選択テクニック(Next-speaker Selection Technique)を用いて次の話し手を選択する。この時、次の話し手に話す義務が生じる。次話者選択テクニックは、以下の4種類が指摘されている。

(a)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、隣接ペアの第一部分である質問部を利用できる。すなわち質問で次の話し手を選ぶ。

(b)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。

(c)「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(Tag Question)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。

(d)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。

② 自己選択規則

現在の話し手が次の話し手を選択せず、他の会話者が自主的に話し手になる。この時、最初に話し始める会話者が発話する権利を得る。

③ 自己継続規則

現在の話し手が次の話し手を選択せず、また他の会話者も自己選択せず、今の話し手は続けてもよいが、その義務が生じない。

エージェントの設計において、会話の場面と会話参加者の参与役割に応じて話者交替パターンを設計することで、研究目的に即した形態の会話へ導ける可能性がある。

2.1.3 修復(Repair)

自然会話において、前節で述べた規則が適用されることによって話者の円滑な交替が実現されるが、交替の誤りや違反も存在しており、それに対処するための修復メカニズム(Repair Mechanism)が指摘されている[7]。

一例を挙げれば、2人の会話者が同時に発話をしていることに気付いた場合、そのうちの1人が途中で話をやめてトラブルを解決する、このような行為が修復の一種にあたる。会話の中に、理解や聞き取りに関する問題をトラブル源(Trouble Source)と呼ぶ。上の例では、2人の会話者の発話の衝突がトラブル源になり、1人の会話者が発話をやめることが修復になる。

修復は、開始(Initiation)と実行(Repair)という二つの要素から成り立つと Schegloff ら[8][9]が述べている。会話参加者がトラブル源に気づいていると、すぐに修復を実行するではな

く、修復の開始によって、トラブル源に対して理解できなかったという認知を提示し、その上で修復の実行を行う。

会話の参加者のうちの誰がトラブルに気づいて修復を開始し、誰が修復を実行するのかについて、開始・実行担い手がトラブル源の話し手自身 (Self) と他の参与者 (Other) に分けられる。要するに修復の開始は、自己開始(Self-initiation)と他者開始(Other-initiation)に分けられ、修復の実行は、自己修復(Self-repair)と他者修復(Other-repair)に分けられる。そして修復を実行する人が修復を開始する人とは限らないため、以下四つのパターン (修復連鎖) がある。

(1) 自己開始修復(Self-initiated Repair)

例[8] :

A: She was givin me all the people that were gone this year. I mean this quarter year.
B: Yeah.

この例では、A が自分の発話‘this year’にトラブルに気づき、‘I mean’で自己開始をし、‘this quarter year’を示すことで自己修復を実行する。

(2) 他者開始修復(Other-initiated repair)

例[8] :

A: Lissena pigeons.
B: Quail, I think.
A: Oh yeh? No that's not quail, that's a pigeon.

この例では、聞き手である B が A の発話‘Lissena pigeons.’に対しトラブルに気づき、‘Quail, I think.’を言って A の誤りを直すことで他者開始をしながら他者修復を実行する。

(3) 他者開始自己修復(Other-initiated self-repair)

例[8] :

A: Have you ever tried a clinic?
B: What?
A: Have you ever tried a clinic?
B: No, I don't want to go to a clinic.

この例では、聞き手である B が A の発話に聞き取れないというトラブルに気づき、‘What?’ で他者開始をする。そして A が言い直して自己修復を実行する。

(4) 自己開始他者修復(Self-initiated other-repair)

例[8]:

A: I can't think of his first name, Watts on, the one that wrote that piece.

B: Dan Watts.

この例では、A が'I can't think of his first name, Watts on'と発話し、名前を思い出せないトラブルを示して自己開始をする。聞き手である A が正しい名前を言い出し、修復を実行する。

円滑なヒューマンエージェントインタラクションを達成するために、エージェントの応答仕組みの質を高める、またはエージェントの言語理解を改善するような工夫がされている。一方、擬人化エージェントに修復メカニズムを設置する手法が、人間らしさを向上させたうえ、エージェントの発言に由来するトラブルを解消できると考えられる。

2.2 擬人化会話エージェント

2.2.1 単体エージェントによる会話

擬人化会話エージェントに関する研究では、ユーザ 1 人がエージェント 1 体と二者会話をするシステムがよく見られる。主に、情報提供を目的とするものが多く研究されている。

例えば、ガイドの分野では、Extempo [10]のようなキャラクターエージェントがあり、ユーザとの会話を通じて、ユーザの要求に応じたウェブページへガイドできる。Max[11]は、博物館のガイドとして、自然との対面コミュニケーションで訪問者と接したり、博物館や展覧会に関する情報を提供したりするエージェントである。WikiTalk[12]のような Wikipedia を知識源として用いたオープンドメインの知識アクセスのための音声対話システムもある。教育や e-learning の分野では、例えば医学教育を目的としたエージェント Adele が開発され、ウェブによる情報を提供しながら、学生の問題を解説する[13]。受講者が人間型ソフトウェアロボットと対話しながら自学自習することを目標とした CAP という e-learning システムもある[14]。

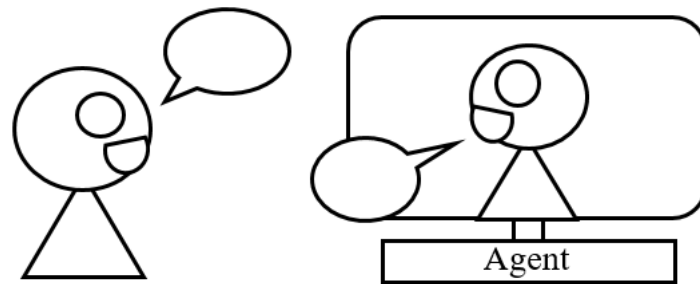


図 3 会話エージェントを用いた二者会話

以上のようなシステムは、ユーザからの質問とシステムによる回答のような単純な二者交替パターンを用いた。図 3 に示すように、エージェントとユーザが相互に話し手と聞き手という役割を交替しながら会話を行い、情報を交換する。

一方、コミュニケーション支援の分野では、エージェントが多人数会話の一員となり、ユーザ同士の会話を支援する研究が見られる。これらのエージェントは、図 4 に示すように、主に傍参与者としてユーザ同士の会話を聞き、特定な場合において割り込んで発話する。

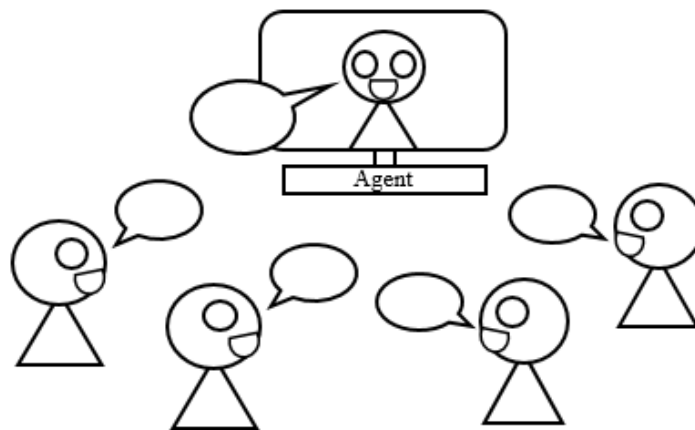


図 4 会話エージェントを用いた多人数会話

例えば、高瀬ら[15]は、会話中の参加者の参与役割と優位性を推定し、特定の参加者の意見を求めて議論を発展させる対話ロボットを作成した。このロボットは、話し手の各参加者への注視時間や相互注視した時間を用いて次話者を推定する。次話者がロボット自体に当たると、ロボットが所定の会話ドメインにおいて、想定質問と応答対応リストを用い

て参加者からの質問に回答する。それに加えて、優位性の低い参加者に対して質問をし、話題を発展させる。

介入直後に参加者間の会話の中で内容語（名詞、動詞、形容詞）が含まれる発話を行った人数を各介入の介入影響度とし、それを評価尺度とした。結果として、参加者の参与役割と優位性を考慮した介入は、そうでない介入に比べて話題を提供し、会話や応答を誘発する可能性が認められた。したがって、参加者の役割と話者交替への考慮が、多人数会話におけるアンバランスの解消につながると言える。

しかし、モデルや会話内容のデータをあらかじめ用意しなければならない点が提案システムの制限であり、任意の環境において適切な介入を行うために会話参加人数やその発話状態、顔向き状態を把握したり、適切な介入内容をリアルタイムで生成したりすることは難しいと述べられている。

2.2.2 複数体エージェントによる会話

単体の会話エージェントより、互いに協調する複数の会話エージェントが人の印象を改善することが検証されている。Sakamotoら[16][17]は、ロボットを広報媒体として用いることを目的とし、2体のロボットを駅に設置し、通行人の反応を調査した。その結果、1体のロボットが話すより、2体のロボットが互いに会話する方が、人が立ち止まって話を聞こうとする傾向が高く、人の印象が良いことが明らかになった。

有本ら[18]は音声認識を用いない、対話感が低い2体の対話ロボットを連携させ、対話感への影響を調べた。複数体ロボットを用いた場合は単体ロボットを用いた場合より、人の印象がよくなったと報告した。具体的には、会話参加者の「会話が難しい」、「ロボットに無視されている」、などの感覚が少なくなったことが分かった。複数のロボットによる発言の分散による影響が理由だと推測されている。また、ユーザが時々「参与はしているが発話が宛てられていない」傍参与者となることも理由の一つだと述べられている。

高橋ら[19]は、高齢者の発話機会の増加を目的とし、護職員やボランティアの代わりに高齢者と対話を行うロボットを開発した。このロボットは、ソーシャルメディア上のコメントを発話文として用い、ユーザである高齢者と自然な文脈で発話でき、ロボット同士の会話も可能である。比較実験として、単体ロボットを用いた二者会話と複数体ロボットを用いた多人数会話が高齢者の発話機会にどのような影響を与えるかを調査した。ロボット同士の話し合いにより、場の雰囲気が盛り上がっているように参加者は感じることができ、また、参加者にポジティブな印象を与えたと報告されている。

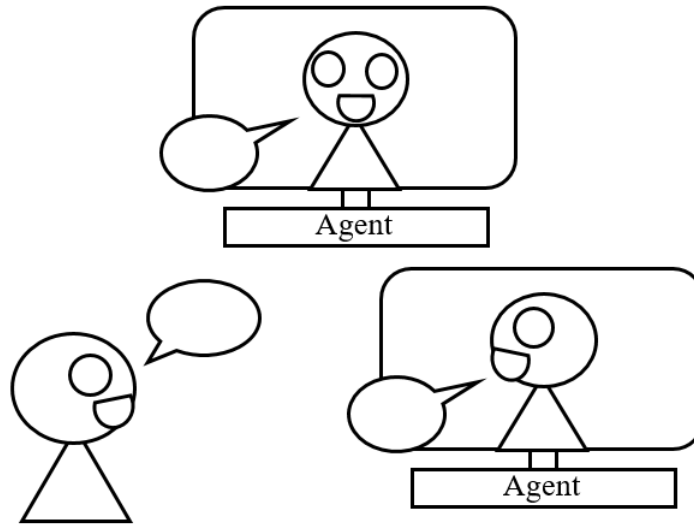


図 5 複数体エージェントを用いた多人数会話

また，複数体会話エージェント同士が話し合いなどのインタラクションを取りつつ，協調して人と会話するようなシステムが開発され，その有効性が指摘されている．図 5 に示すように，藤堂ら[20]は単体エージェントを用いた二者対話システムを比較的簡単に拡張し，複数体エージェントを用いた三者対話システムを構築した．このシステムは雑談を目的とし，会話ドメインと三者対話の状態遷移をあらかじめ設定した．例えば，「北海道旅行と沖縄旅行」のドメインにおいて，エージェント L が北海道旅行について勧め，エージェント R が沖縄旅行について勧める．

エージェントの交替パターンとして，まずはエージェント L から北海道について質問し，ユーザの返事に対して，続いて何回も質問する．そして「だよね」などの付加疑問 (Tag question) を使ってコメントした後，2 体のエージェントが交替し，エージェント R がエージェント L と同じような順番で，沖縄についてユーザと会話する．この繰り返りで対話が進んでいく．

実験結果として，同じ対話ドメインであっても，二者対話と三者対話で，ユーザの印象の違いが見られ，二者対話システムで不十分であった雑談らしさが，十分に改善されたと述べられている．また，対話エージェントが交替することで，ユーザの音声認識誤りのストレスを軽減させている評価も見られた．

Iio ら[21]は 2 台のロボットを用いて，人の発話を誤認識し，矛盾している応答をしても，それを感じにくくさせる話者交替パターンを提案した．具体的には，ロボット A の質問に対する人の返答に対し，ロボット A が「うん…」などの感嘆詞 (Interjection) を使って曖昧な応答をし，有意に発話を終了して発話権を譲る．ロボット B が人の返答から認識された単語に対してコメントをしたり共感的な応答をする．この話者交替パターンによって，

誤認識が生じてエージェントが別の単語に対してコメントを行っても、ユーザが適当なあいづちを打つことで対話が進むようになっている。参加者は対話の矛盾が減少し、整合性が維持されているように感じた。

この結果から、感嘆詞を用いた話者交替パターンによって、社会的文脈の構築に成功したことが示唆された。さらに、そのような矛盾の減少は参加者にロボットと深く話し合ったという印象を与えたことも考えられる。

以上述べた研究から、複数体エージェントからなる会話システムは、多様な会話形態があるゆえ、単体システムのない効果に結び付くことが分かった。具体的には、

(1) 参与役割の変化による印象の向上

複数のエージェントによる他人数会話では、ユーザが時々傍参与者(side-participant)となってエージェント同士での話し合いを見ることがある。これが会話感の向上と場の雰囲気作りにつながると推測されている。

(2) エージェント間の交替による矛盾の軽減

エージェント間の交替に伴う話題の転換・分裂や立場の切り替えなどが、単一な話題における齟齬を防ぐ効果がある。2体のエージェントは二つの独立な存在であるため、事前に定義された応答内容の制限や音声認識の誤認識などの原因で矛盾な文脈が生じても、話題・立場の両立によって緩和される可能性がある。

第二言語会話支援場面の課題解決において、これらの知見も適用され得るだろうと考えられる。

2.3 本研究の位置づけ

本研究は、第二言語会話支援を目的とし、NNSの発話権取得の困難さに着目する。対面会話の基本的な規則をいかに擬人化会話エージェントに適用し、それがいかにNNSの発話権取得を支援できるかを研究している。図6に本研究の位置づけを示す。

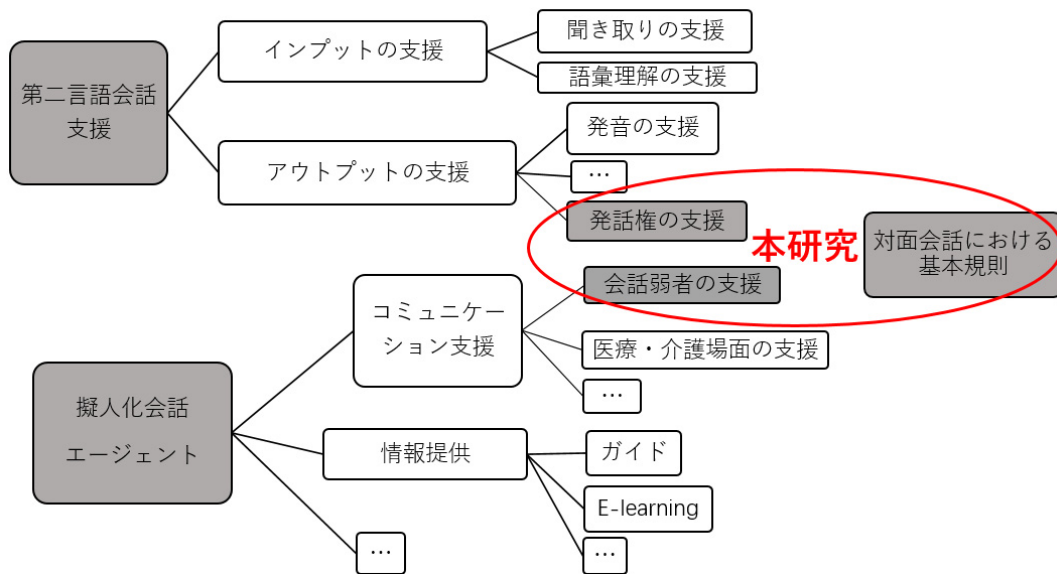


図 6 本研究の位置づけ

第3章 NNS 支援エージェントの概要と評価実験

我々は、NNS 支援エージェント[4]を提案している。NNS 支援エージェントが、NS と NNS の会話行動にどのような影響を与えるか、その介入の効率がどの程度かを探るために、会話実験を行った。本実験は、筑波大学図書館情報メディア系倫理審査委員会の承認(通知番号第 29-141 号)を得た。本章では、NNS 支援エージェントの概要とその実験の詳細を述べる。

3.1 NNS 支援エージェントの概要

NNS 支援エージェントは、話者交替規則と現話者による次話者選択テクニックを利用し NNS に発話権を渡すことを目標とした。図 7 が示したように、3D キャラクターが第二言語会話の参加者となり、音声検出によって NS と NNS の音声を常時検出しでき、音声合成によって発話できる。

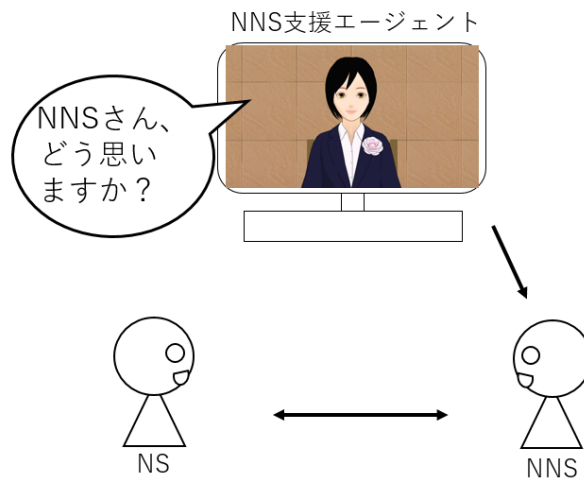


図 7 NNS 支援エージェントを用いた第二言語会話

エージェントの発話内容として、発話権を NNS に渡すために、2 種類の次話者選択テクニックに基づいて 4 つの質問文を設計した (表 1) .

表 1 エージェントの発話内容

No.	次話者選択テクニック	発話内容
1	(a)	○さん、どう思いますか?
2	(a)	○さん、どうでしょうか?
3	(c)	○さん、そう思わないの?
4	(c)	○さんもそう思いますよね.

エージェントの介入タイミングとして、会話中に他者選択規則も自己選択規則使われなく、自己継続規則も使われない状態が続くと、エージェントが介入するように設計した. システム上の実現として、図 8 に示すように、NS と NNS の両方が 2 秒の沈黙に入る [22]と、エージェントが発話内容からランダムに一つを選んで発話する.

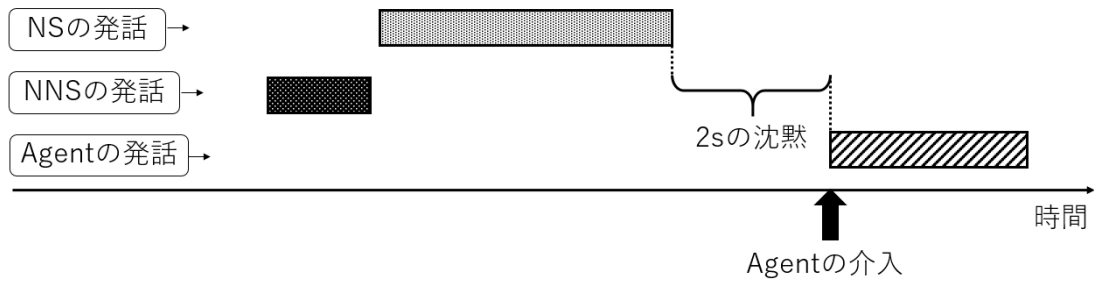


図 8 エージェントの介入タイミング

3.2 参加者と条件

実験参加者は、初対面かつ同じ性別[23]の日本語母語話者 1人と日本語非母語話者 1人を 1ペアとし、男性・女性それぞれ 9ペアであり、計 18 ペアである。すべての参加者は大学生または大学院生であった。

次の 3つの条件で参加者内計画の実験を行った。すなわち、各ペアの参加者に会話を 3回行わせた。各条件への参加順序はカウンターバランスをとった。実験の風景を図 9に示す。

(1) エージェント発話条件

NNS 支援エージェントが参加者の会話に介入する。

(2) エージェント非発話条件

エージェントが参加者の会話に介入しないほかは(1)と同一とした。

(3) 非エージェント条件

参加者の二人が会話を行う。エージェントがいないほかは(1)(2)と同一とした。

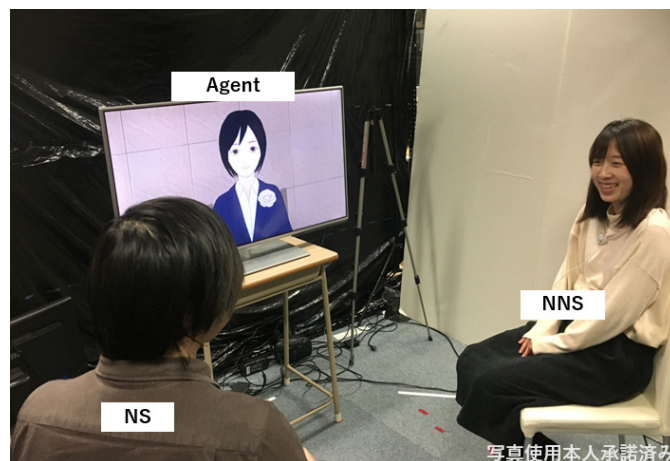


図 9 実験の風景

3.3 会話タスク

会話タスクは、一つの話題についての8分間の会話にした。話題は、会話開始前に筆者から提示される話題の中から参加者の二人が選択することで決定した。(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)。

3.4 環境と手順

実験は筑波大学春日エリアの7C204研究室で実施され、異文化交流の場面を想定するため、雑談スペースを構築した。環境騒音は40dbであり、参加者間の対話距離を1mとした(図10)。

ソフトウェア環境として、OSがWindows7であり、制御スクリプトがpython3.6で実行される。デバイスとして、エージェントを映すに使うディスプレイはPANASONIC TH-L42E60(サイズ:幅957mm×高さ619mm×奥行217mm,画素数:1920×1080)である。音声検出に使用しているマイクはSteelSeries Siberia Microphone white 50002(周波数:80-15000Hz,感度:-38dB)である。録画に使うカメラはSANYO Xacti DMX-HD2000(解像度:1920×1080)であり、全景を撮るためのワイドコンバージョンレンズはSANYO VCP-L07W(0.7倍)である。

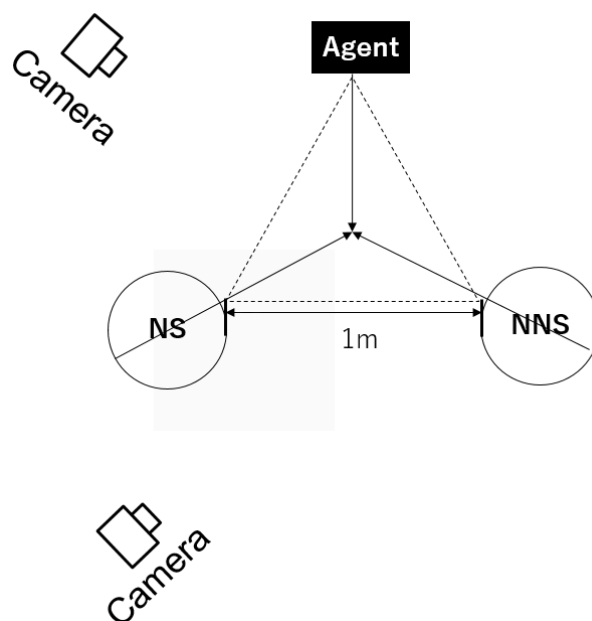


図10 実験の配置

手順としては、会話開始前に、まずマイクの装着と動作確認を行った。そして話題の選択肢が書かれた用紙を参加者に提示し、一つ的话题を話し合いで決定してもらった。次に、ディスプレイに投影された等身大サイズの女性 3D キャラクター（実験参加者と同席しているように設定したため、上半身のみ見える）の前で、二人の参加者がそれぞれ椅子に座り、会話を行ってもらった。各ペアに 8 分間の会話を 3 回行わせ、毎回違う話題にした。会話が終了する度に、質問紙を記入してもらった。

3.5 データ収集

3.5.1 会話行動

エージェントと参加者との会話は、2 台のビデオカメラにより記録した。エージェントを映すモニタと参加者の全身が見えるように撮影した。

3.5.2 質問紙調査

コミュニケーションの状況及び会話への印象を評価するため、異文化コミュニケーションや会話分析でよく使われているコミュニケーションの自然さ、発話しやすさなどの項目を利用した[24][25][26][27]。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、エージェントが NNS の発話を促すことができたかどうかなどの項目を利用した。全ての項目は 7 段階（1～7 点）で評価してもらった。

3.6 データ処理

3.6.1 データ量

分析には各条件で挨拶などの多い最初の 1 分を除いた。一人の参加者において、分析に使ったデータ量は 3 条件×7 分=21 分であり、合計データ量は 3 条件×7 分×18 ペア×2 話者 = 12 時間 36 分 = 756 分 = 45360 秒であった。

3.6.2 発話区間のラベリング

図 11 にラベリングの様子を示す。録画したビデオはアノテーションツール ELAN[28]のセグメンテーション機能を用いて、最小無音区間 300 ミリ秒の条件で発話区間を自動分割した（間終止単位、IPU）。

次に筆者により発話区間ごとにその発話者（NS,NNS,エージェント）を判別し、人手でラベリングした。ラベリングする際に、発話区間の確認をし、自動分割がうまくできない箇所を人手で修正した。また、笑い声や咳、溜息などの非言語有音区間は発話区間として扱わなかった。

ラベリングが完了した後に、発話者ごとに図 11 (B) のような発話区間のリストを EXCEL ファイルに書き込み、発話者ごとの発話区間の数を数え、発話者ごとの発話合計時間を計算した。

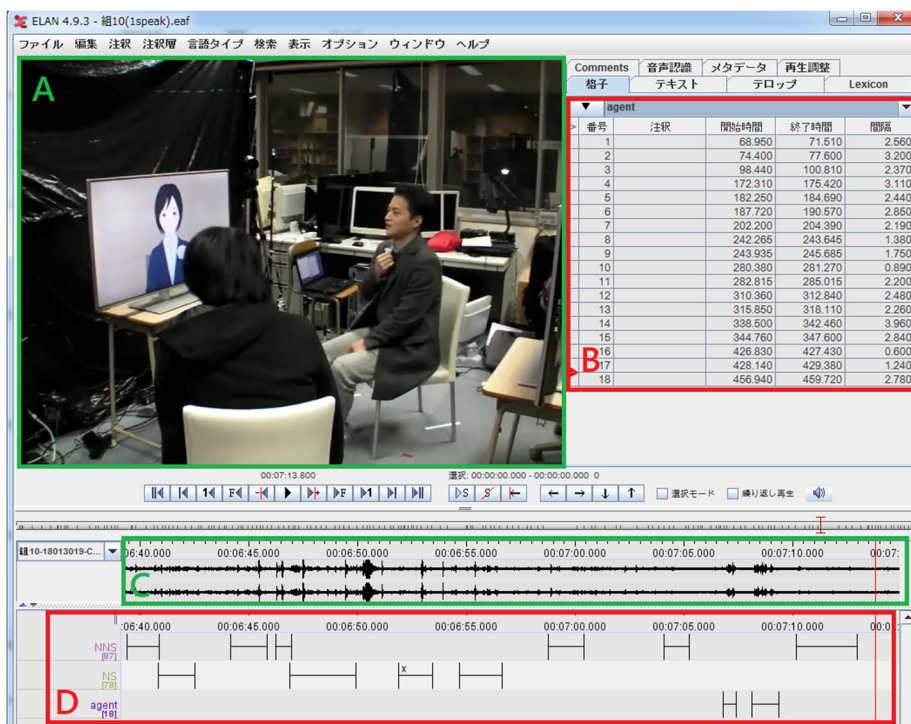


図 11 発話区間ラベリングの例：会話の映像データ (A)、会話音声データの波形 (C)、NNS・NS・Agent それぞれの発話区間のラベリング (D) ラベリングした発話区間のリスト (B)

3.6.3 ターン（発話権）の判別とラベリング

発話区間のラベリングをもとに、筆者が会話の内容を聞きながら NNS・NS・Agent それぞれのターン（発話権）を判別し、図 12 のようにラベリングした。

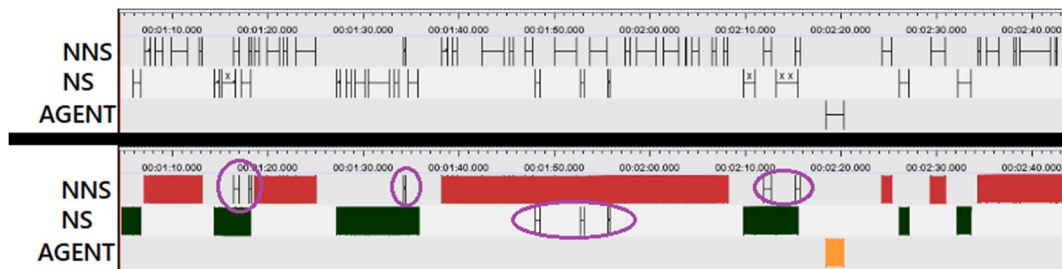


図 12 ターンラベリングの例：ELAN で作成した発話区間のラベリング（上），画像編集ソフトで作成したターンのラベリング（下），NNS・NS・Agent それぞれのターン（赤色・緑色・黄色の部分），あいづち（紫色の丸で囲まれている部分）

画像編集ソフト（Adobe Photoshop CS 5[29]）のペンツールを使い，一人の参加者において，ターンを取るからほかの参加者が次のターンを取るまでの間にある発話区間の和集合を一つの色でマーキングした。

一つの発話に対し，発話者がこの発話によってターンを取ったかどうかを判断するには，この発話があいづちであるかどうかによって決定した。あいづちは，「話し手が発話権を行使している間に聞き手が送る短い表現」と定義され[30]，聞き手が発話権を取ろうとしない，話し手の話をきいていることを示す発話である。あいづちの認定基準[31]を参考に，以下の 6 種類の発話をあいづちとして扱った。

- 「ああ」「うん」「ええ」「はい」などの応答系感動詞。
- 「あっ」「えっ」「おっ」などの感情表出系感動詞。
- 「なるほど」「確かに」「そう（ですね）」などの語彙的応答。
- 相手の発話の繰り返し（疑問型上昇調を除く）。
- 相手の発話の補完。
- 「おもしろい（な）」「すごい」「こわ」などの評価応答。

一つの発話が以上の 6 種類のいずれかに当てはまったら，あいづち表現として認定し，話者がターンを取らないと判断した。従って，この発話に色を付けなかった。ターンのラベリングが完了した後に，NNS・NS のあいづち数とターンの交替する回数を数えた。

3.6.4 介入分類のラベリング

エージェントの介入効率を調べるため，エージェント発話条件の映像データを観察し，介入の分類を行った。エージェントによる介入は合計 75 回であった。そのうち，会話途中で状況を理解できずに，実験実施者に問いかけたケース（計 5 回）を除き，70 回を分類の対象とした。一回の介入において，介入時点の約半分前後にわたる会話内容を書き出し

た. そのコンテキストと介入参加者の言語反応に基づき, 介入を以下のように分類してラベリングした.

(a) **直接有効介入 (DS, Direct Success)** : NNS エージェントの介入が NNS の発話を促すことができた介入を有効介入と定義する. 特に, エージェントから NNS に発話権を直接に渡した場合を直接有効介入 (DS) と定義する. 例を表 2 に示す (会話例にある「○」は, NNS の苗字を指す. 以下同じ).

表 2 DS の会話例

ペア番号 : 4	第 2 分 10 秒	「部活」
NS : ほかの学年でも仲良くなることができますし...		
NNS : あー		
NS : ...はい.		
AGENT : ○さんもそう思いますよね?		
NNS : はい.		

(b) **間接有効介入 (IS, Indirect Success)** : エージェント介入の直後に, NS がエージェントの質問を繰り返して, 従って発話権を NNS に渡した場合を間接有効介入 (IS) と定義する. 例を表 3 に示す.

表 3 IS の会話例

ペア番号 : 11	第 2 分 8 秒	「ワールドカップ」
NS : 今年はどうなんですか? 中国は強いですか? サッカー.		
NNS : んん...つよ.....		
AGENT : ○さん, どうでしょうか?		
NS : わお, どうでしょうか?		
NNS : そうですね. 中国は...強い...ではないです.		

(c) **反応あり介入 (RF, Responded Failure)** : エージェントの介入によって, 発話権を NNS に渡すことができなかったが, エージェントの介入をきっかけとして, 参加者が言語反応 (例えば評価応答, 感情表出系感動詞などの発話, またはエージェントについての話し合いなど) をした場合を RF と定義する. 例を表 4 に示す.

表 4 RF の会話例

ペア番号 : 13	第 5 分 44 秒	「野菜」
NS :	なんか、中国から日本に来て、初めて見た野菜	
	とかありますか？	
NNS :	初めて見た野菜.....	
AGENT :	○さん、どうでしょうか？	
NS :	えへへ？	
NNS :	うふふ(笑).	
NS :	支援してくれる.	
NNS :	すごいですね. 初めて見た... うーん... 思い出	
	せないですね.	

(d) 反応なし介入 (NF, Non-responded Failure) : エージェントの介入によって、発話権を NNS に渡すことができず、その介入が参加者に無視された (言語反応がない) 場合を NF と定義する. 例を表 5 に示す.

表 5 NF の会話例

ペア番号 : 1	第 2 分 15 秒	「旅行」
NS :	あの、南のほうはあんまりいかないですか？旅行に.	
NNS :	なんか、京都と大阪が、今年の...の...	
AGENT :	○さんもそう思いますよね？	
NNS :	十、十です.	
NS :	十月.	
NNS :	そう.	
NS :	に、旅行に行って.	
NNS :	そうです.	

介入分類のラベリングが完了した後に、DS・IS・RF・NF それぞれの回数を数えた.

3.6.5 介入前後の会話内容の分析

介入に伴う発話権譲渡によって、会話の成り行きがどのような変化が起きたか、特に発話権譲渡が失敗した場合、会話がどのように進行し続けたかを調べるために、RF と NF を対象とし、介入前後の参加者の言語反応を分析した.

NF には、介入直後に言語反応がないため、介入直前の会話内容から、無反応の結果をもたらすことが可能な表れを探った。

RF には、参加者がエージェントの質問に反応をしたが回答をしなかった、いわゆるエージェントの質問への受理ができなかった。Clark ら[32]の共有基盤 (Common Ground) 理論によると、エージェントの発話がトラブルとなったと考えられている。したがって、Schegloff ら[8]の定義を判断基準とし、エージェント介入後の修復行為を分析した。具体的に、エージェントの発話がトラブル源になった場合、エージェントの自己修復が設定されていないため、聞き手 (NS と NNS) による他者修復のみが可能である。それ故に、エージェントの介入後に、他者開始について分析した。他者開始の発生位置については、Schegloff らが「トラブル源の出現したターンの直後」だと指摘した。他者開始の形式については、鈴木[33]が Schegloff らの指摘した 7 種類に従い、日本語における他者開始形式を表 6 のように整理した。

表 6 日本語における他者開始形式 (鈴木[33])

No.	フォーマット	発話内容
1	なし	え？ん？はい？あ？は？何？
2	疑問詞+上昇イントネーション	誰？いつ？どこ？
3	部分繰返し+疑問詞	おばあちゃんが何？
4	部分繰返し+引用形式「って」 (+疑問詞)	そちらって：/どこ？それ。
5	なし	どういうこと？/よくわからない
6	全部/部分繰返し (+動詞句に「の」を付加)	今日？/日比谷線？/行ったの？
7	i. Xじゃない？ ii. 部分繰返し+引用形式「って」+理解候補	藤木直人じゃなくて？ もうちょっとって1時間ぐらい？

以上に基づいて、エージェントの介入直後に他者開始形式の有無を判断してラベリングした。エージェントの介入直後に他者開始が発生した一例を表 7 に示す。

表 7 介入直後に他者開始が発生した例

ペア番号：14	第3分35秒	「牛肉」
---------	--------	------

NS：なんか日本で焼肉屋さんとかお肉屋さんとかでもも肉って売ってないですよ。

NNS：ああ、うん、そうそう。

NS：それは堅いからですか？

NNS：たぶん...あの...うーん...

AGENT：○さんもそう思いますよね。

NNS：どっちが言いました？（他者開始）

NS：聞き取れなかったです。

NNS：わかんない、ふふ。

NS：なんだろ。

AGENT：○さん、どうでしょうか？

NNS：何がどうでしょうか？（他者開始）

NS：わかんない。

3.6.6 内容語のラベリング

3.6.4 で述べた 4 種類の介入は、介入直後の会話にどの程度の影響を与えたかを定量的に評価するために、介入影響度[15]という尺度を使った。それを計算するために、エージェントの介入へ音声反応がある介入（DS・IS・RF）を対象とし、介入直後のターンとその次のターンの会話内容に、内容語（名詞、動詞、形容詞）の有無をラベリングした。

第4章 実験結果と考察

4.1 発話頻度

発話区間の数を発話数とし、1分あたりの発話数を発話頻度とした（図 13）。

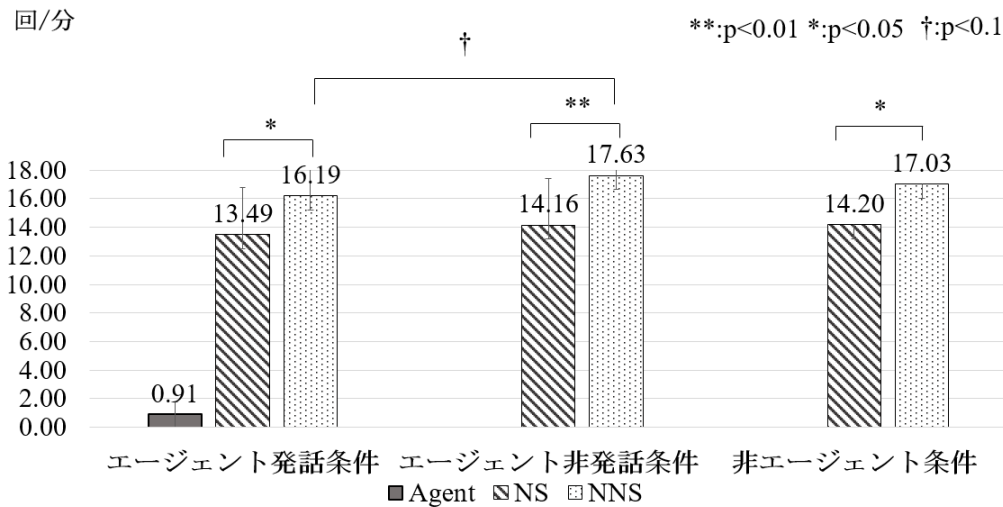


図 13 発話頻度(N = 18)

エージェント発話条件で NS が $M = 13.49$, $SD = 3.67$, NNS が $M = 16.19$, $SD = 4.04$, エージェントが $M = 0.91$, $SD = 0.89$. エージェント非発話条件で NS が $M = 14.16$, $SD = 3.31$, NNS が $M = 17.63$, $SD = 3.25$, 非エージェント条件で NS が $M = 14.20$, $SD = 4.11$, NNS が $M = 17.03$, $SD = 3.57$ である。

条件間で比較した結果, NNS はエージェント発話条件の方でエージェント非発話条件より発話頻度が減る傾向が見られた (一元配置分散分析, $F(2,34)=.989, SE=5.231, p<0.1$) .

話者間で比較した結果, いずれの条件でも NNS は NS より発話数が有意に多かった (エージェント発話条件: t 検定, $t(34)=2.098, p=.043$; エージェント非発話条件: t 検定, $t(34)=3.171, p<0.01$; 非エージェント条件, t 検定, $t(34)=2.209, p<0.05$) .

この結果から, いずれの条件においても NNS は NS より頻繁に発話したことが分かったが, エージェントの介入による影響が見られなかった. 言語能力の違いがあるため, NS が流ちょうに発話できるが, NNS の一連の話が短い発話に区切られ, また, 言い直しが多い可能性があると考えられる.

4.2 発話量

一回の会話につき, NS・NNS それぞれの発話区間の継続時間の和をこの回の合計発話量とし, 1 分あたりの発話量を図 14 に示す.

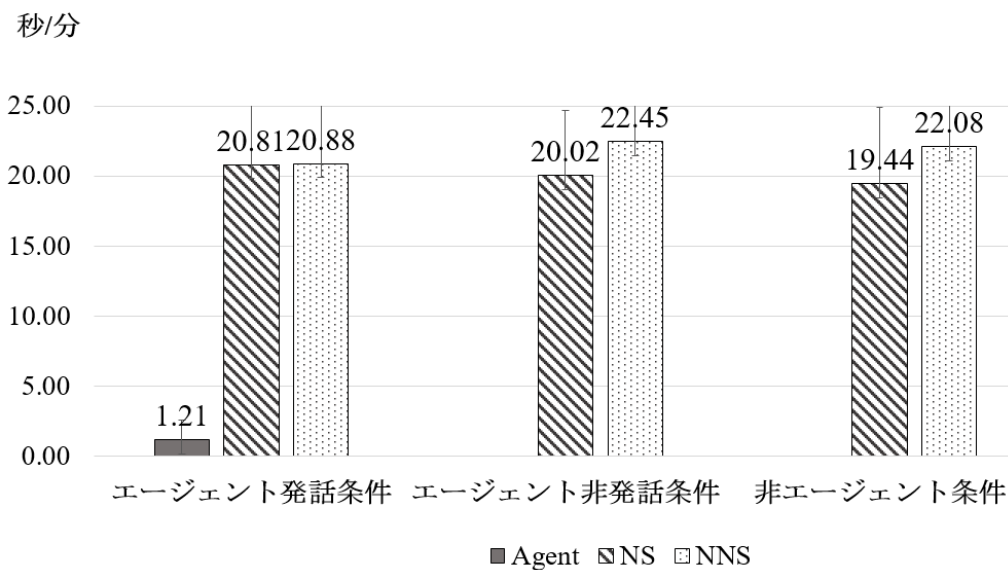


図 14 発話量(N = 18)

エージェント発話条件で NS が $M = 20.81$, $SD = 7.24$, NNS が $M = 20.88$, $SD = 5.42$, エージェントが $M = 1.21$, $SD = 1.41$. エージェント非発話条件で NS が $M = 20.02$, $SD = 4.66$, NNS が $M = 22.45$, $SD = 5.72$, 非エージェント条件で NS が $M = 19.44$, $SD = 5.45$, NNS が $M = 22.08$, $SD = 6.72$ である.

条件間で比較した結果, 有意差が見られなかった (一元配置分散分析) .

話者間で比較した結果, いずれの条件でも有意差が見られなかった (t 検定) .

この結果から, いずれの条件においても NNS と NS の発話にかかった時間が同じであることが分かった. そしてエージェントの介入による影響が見られなかった.

4.3 発話長

1 回の発話にかかる時間の長さを発話長とした. まず 1 人の参加者が 1 回の会話における平均発話長を計算した. それを 1 条件のサンプルとし, 18 人の NS・NNS が 1 つの条件における平均発話長の平均値を計算した (図 15) .

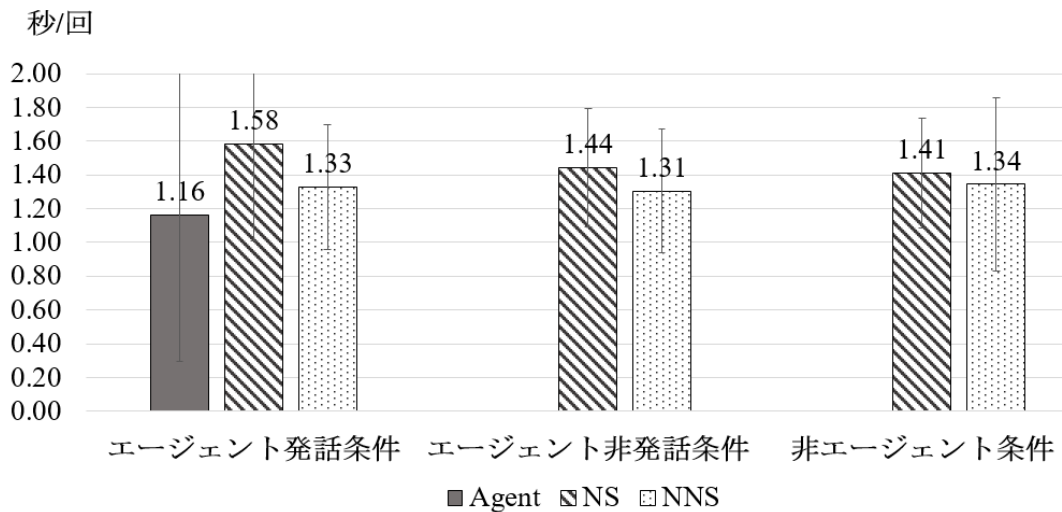


図 15 発話長(N=18)

エージェント発話条件で NS が $M = 1.58$, $SD = 0.58$, NNS が $M = 1.33$, $SD = 0.37$, エージェントが $M = 1.16$, $SD = 0.87$. エージェント非発話条件で NS が $M = 1.44$, $SD = 0.35$, NNS が $M = 1.31$, $SD = 0.37$, 非エージェント条件で NS が $M = 1.41$, $SD = 0.33$, NNS が $M = 1.34$, $SD = 0.51$ である.

条件間で比較した結果, 有意差が見られなかった (NS の発話長 : Friedman 検定, NNS の発話長 : 一元配置分散分析) .

話者間で比較した結果, いずれの条件でも有意差が見られなかった (エージェント発話条件 : Mann-Whitney 検定 ; エージェント非発話条件 : t 検定 ; 非エージェント条件, t 検定) .

この結果から, いずれの条件においても NNS と NS の発話長が同じであることが分かった. そしてエージェントの介入による影響が見られなかった.

4.4 あいづち率

参加者たちが会話においてどれくらい聞き手・話し手の参与役割を担当するかを調べるため, 発話数に対するあいづちの比率 (あいづち数 / 発話数) を計算した (図 16) .

エージェント発話条件で NS が $M = 23.15\%$, $SD = 0.13$, NNS が $M = 19.29\%$, $SD = 0.13$. エージェント非発話条件で NS が $M = 21.53\%$, $SD = 0.13$, NNS が $M = 15.17\%$, $SD = 0.08$, 非エージェント条件で NS が $M = 22.95\%$, $SD = 0.11$, NNS が $M = 15.10\%$, $SD = 0.09$ であった.

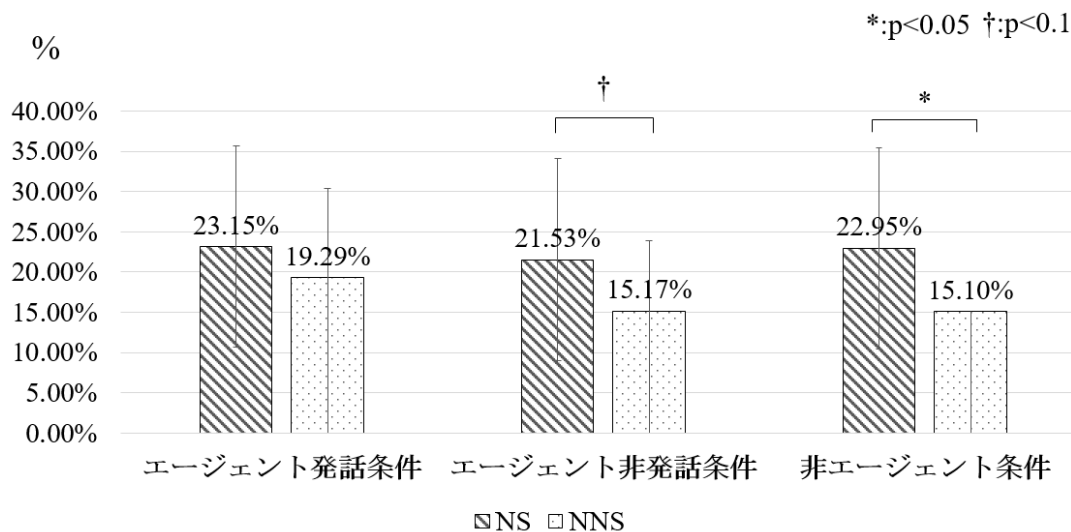


図 16 あいづち率 (N=18)

条件間で比較した結果、有意差が見られなかった (Friedman 検定)。

話者間で比較した結果、エージェント非発話条件も非エージェント条件も、NS は NNS よりあいづちが多い傾向が見られた。(エージェント非発話条件 : Mann-Whitney 検定, $u=105.5, w=276.5, p<0.1$; 非エージェント条件 : Mann-Whitney 検定, $u=87.0, w=258.0, p<0.05$)。

この結果から、エージェントの介入を伴う会話は参加者間のあいづち行動の差を減少させ、参与役割担当を平準化させるという傾向があると考えられる。

4.5 発話均衡度

より均等に会話に貢献することが、参加者たちがより親和的で活発な議論を経験することの表現であると指摘されている[34]。前節で述べたように、NS と NNS のあいづち率の差が縮まる一方で、あいづち以外の発話のバランスがどのように変わったかを調べるため、発話均衡度を計算した。あいづちを除き、発話数の少ない参加者の発話数を、発話数の多い参加者の発話数で割った値を発話均衡度とした (図 17)。エージェント発話条件はほかの条件より割合が高かった。

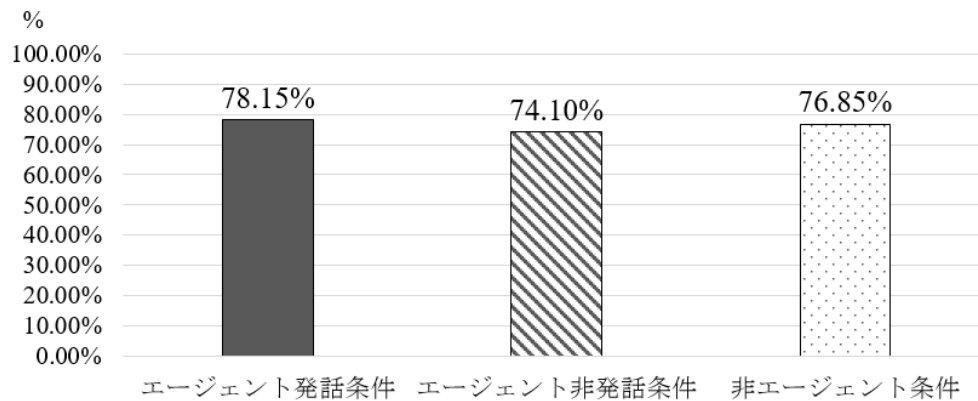


図 17 発話均衡度 (N=18)

4.6 話者交替頻度

ターン（発話権）の移り変わりの回数を話者交替数とし、1分あたりの話者交替数を話者交替頻度とした（図 18）。

エージェント発話条件で $M = 7.63$, $SD = 2.12$ 。エージェント非発話条件で $M = 7.44$, $SD = 3.07$ 、非エージェント条件で $M = 6.32$, $SD = 2.33$ であった。

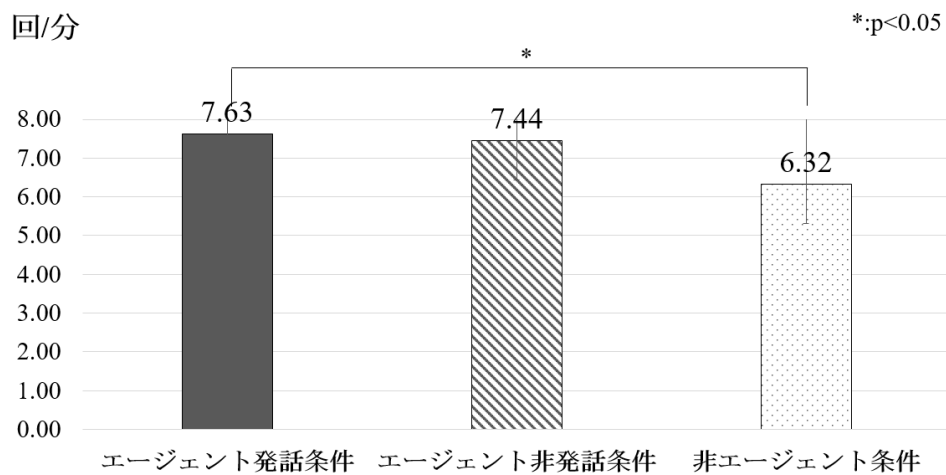


図 18 話者交替頻度 (N=18)

条件間で比較した結果、エージェント発話条件は非エージェント条件より話者交替頻度が有意に多かった（Friedman 検定, $\chi^2[2]=7.000$, $p<0.05$ ）。

この結果から、エージェントがいると、会話中のやり取りが増加し、第二言語会話が活発になる傾向が見られた。

また、統計的有意差が認められなかったが、エージェント非発話条件は、非エージェント条件より話者交替頻度が多くなるような様子も見られた。このことについて、発話の有無にかかわらず、具現化されたエージェントの社会的存在が会話に有益な影響を及ぼしたと思われる。会話エージェントの具現化による効果を調査する先行研究[34]でも同じ結論に達した。

4.7 介入の分類

表 8 に、介入の分類結果を示した。全 70 回の介入のうち、22 回が有効で、48 回が有効でなかった。有効介入のうち、直接に NNS へ発話権を渡したのは 12 回 (DS)，NS 経由で NNS へ発話権を渡したのは 10 回 (IS) であった。有効でない介入のうち、28 回が反応あり (RF)，20 回が反応なし (NF) であった。

表 8 エージェント介入の分類結果

分類	DS	IS	RF	NF
回数	12	10	28	20
%	17.14	14.29	40.00	28.57

この結果から見ると、提案エージェントの有効率は 31%程度であった。本エージェントの介入は極めて単純であるが、一定の効果が見られた。また、IS が 14.29% を占めた。表 3 の会話例のように、エージェントの介入をきっかけとして、NS が意図的に NNS に配慮し、意見を求めるように見えるケースがあるが、NS のパーソナリティによるところも大きいと推測される。

しかし、28.57%の介入では、参加者がエージェントに声かけられたにも関わらず、介入前の話を続けた。40.00%の介入 (RF) では、参加者がエージェントの発話に対し「すごいですね」、「びっくりした」、「うん」のような反応をするだけで、その問いかけに回答しなかった。エージェントが提案手法によって自己選択し、NNS を次話者として選択をしたが、その話者交替がうまくできなかった。それゆえに、エージェントの介入に理解や聞き取りに関する問題が生じた、いわゆるエージェントの発話がトラブル源(Trouble Source)になったことが原因だと推測される。次の節で詳しく考察した結果を述べる。

4.8 介入前後の会話内容の分析

全 28 回の RF 介入のうち、9 回に修復行為が見られた。具体的に、エージェントの発話（トラブル源）に対し、参加者により理解できなかったという認知を示す発話（他者開始）が行われたが、そのあとに修復の実行に当たる発話が見られなかった。

全 20 回の NF 介入のうち、11 回が表 5 の例のように、エージェントは NNS が考え込んだ途中に介入した。そのため、NNS は声かけられたにも関わらず、介入前の話を続けただけであった。NNS が考え込んだ時に、沈黙はしているが、まだ発話権を保ち続けたい傾向、すなわち自己継続の傾向があると思われる。このことから、「一定の沈黙時間が経ったかどうか」という介入タイミングの判別条件の不足が見られた。

なお、RF と NF のうち、エージェントと参加者の発話が衝突（Overlap）したケースが 26 回見られた。サクスら[7]によると、衝突のない話者交替は軽い衝突に伴った話者交替と同じく一般的であり、それらは話者交替の大部分を占める。すなわち、衝突の発生は一般的で短い。しかし、エージェントの設計として、話始めると途中で止まることができないため、長時間の衝突を避けることができない。それが RF と NF の発生につながると推測される。

4.9 介入影響度

4 種類の介入は、介入直後の会話にどの程度の影響を与えたかを定量的に評価するために、介入影響度を計算した。

前章で述べたように、介入直後のターンとその次のターンを対象とし、内容語（名詞、動詞、形容詞）があるかどうかをラベリングした。それに基づき、名詞、動詞、形容詞の内のいずれかが含まれていたら、そのターンの影響度を 1 とし、いずれも含まれていなかったら、そのターンの影響度を 0 とした。したがって、一回の介入の影響度が 0 から 2 になる。

例を挙げると、表 4 の会話例では、NS と NNS がエージェントの質問に対して「えへへ」「うふふ」などのあいづち（ターンとして認めない）をした後、それぞれ「支援してくれる」、「すごいですね」のような反応をした。それらのターンには動詞と形容詞があるため、この回の介入影響度を 2 とした。

介入影響度の結果を図 19 に示す。

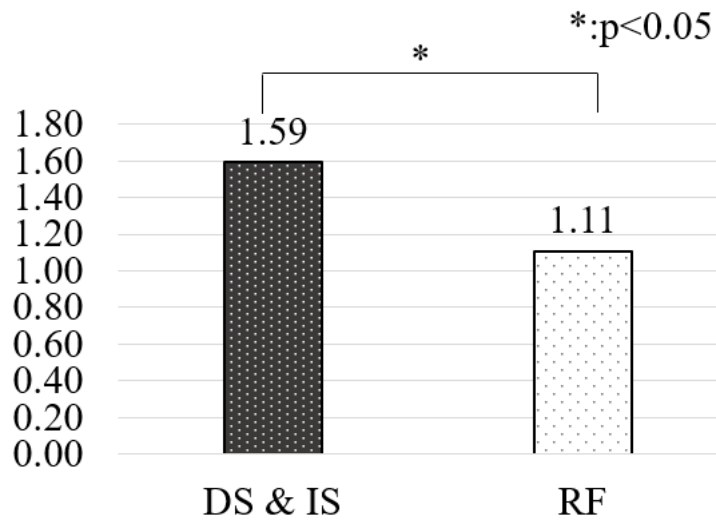


図 19 介入影響度(DS と IS: N = 22, RF: N = 28)

有効介入 (DS と IS) の平均介入影響度が $M = 1.59$, $SD = 0.67$. 反応あり介入 (RF) が $M = 1.11$, $SD = 0.79$ であった. Mann-Whitney 検定で有意差が認められた ($u = 201.5$, $w = 607.5$, $p < 0.05$). このことから, 同じくエージェントの介入が気づかれて言語反応があっても, 発話権譲渡が成功した介入が, 参加者の会話行動に対して比較的大きい影響を与えたことが分かった.

4.10 質問紙調査

コミュニケーションの状況及び会話全体への印象を評価する項目を図 20 に示す.

「相手と会話していて疲れた」の項目において, NS は三条件間で差がなさそうだが, NS はエージェント発話条件もエージェント非発話条件も非エージェント条件より低い点数をした. そのほかの項目において, 5~6 点程度の評価をもらったが, 三条件間で大きな差が見られなかった.

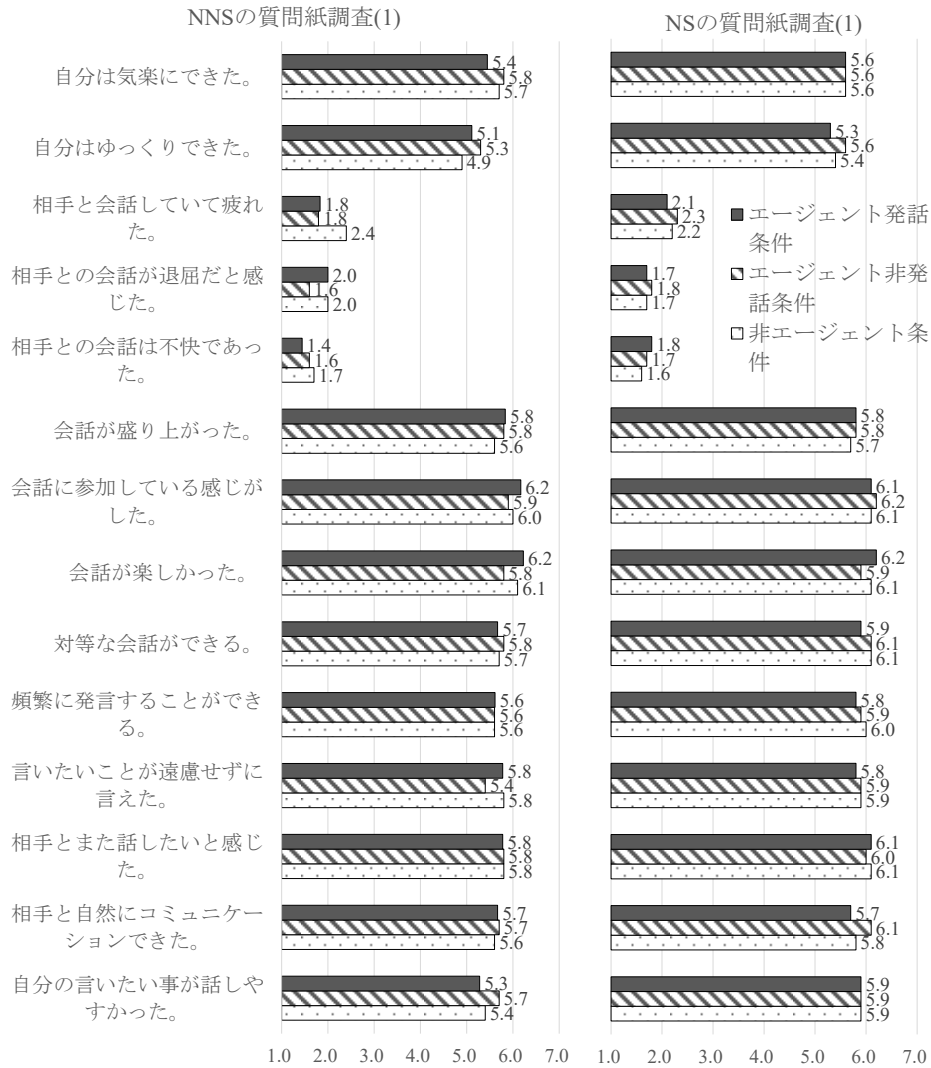


図 20 質問紙調査(1/2) (N=18)

エージェントを評価する項目を図 21 に示す。

図 21 によると、エージェントについて NNS は NS より良い評価を与えたように見えた。NNS は、エージェントから声を掛けられたり、名前を呼ばれたりと能動的な関わりを受けたことが原因かもしれない。しかし、NNS でもおよそ「どちらでもない」との評価であり、総じてエージェントの評価は低かった。本エージェントの介入が極めて単純かつ種類の少ないことが、この結果につながったと考えられる。

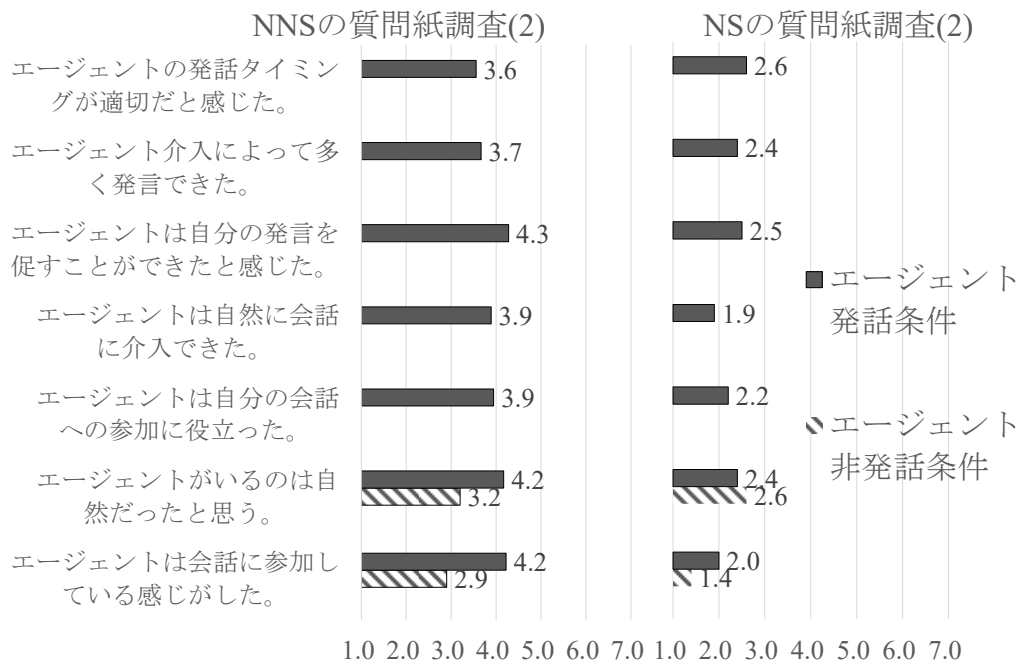


図 21 質問紙調査(2/2) (N=18)

4.11 結論

本章では、話者交替規則に基づく NNS を支援する会話エージェントを用いた実験の結果と考察について報告した。エージェントの介入効率について、極めて単純な介入様式のエージェントであっても会話への第 3 者として加わることで、31%の有効介入ができた。そして有効介入のほうは、有効でない介入より会話への影響度が有意に高かった。NNS への発話権譲渡が成功した場合は失敗した場合より参加者の会話行動に対して比較的大きい影響があることが分かった。次に、この介入効率と影響度を前提として会話の変化を述べる。

発話量と発話長の結果によると、エージェントの存在と介入が NS・NNS それぞれの発話行為に影響がないように見えた。しかし、あいづち率の結果によると、発話の種類割合（あいづちであるかどうか）に変化があるため、エージェントの介入が参与役割担当のを平準化させる傾向が見られた。また、話者交替頻度によると、会話中のやり取りが有意に多くなり、会話が活発になることがわかった。

介入前後の会話内容によると、介入タイミングの判別条件の不足と衝突の回避策を欠いていることが分かった。また、エージェントの介入に由来する不理解の表明（他者開始）に対し、修復の実行が少ないことが分かった。

主観評価について、総じて会話感の評価は一般的であり、エージェントの評価は低かった。そしてNNSとNSの評価の違いがあり、特にエージェントについてNNSはNSより良い評価を与えた。

以上を踏まえ、エージェントの介入効率の向上と参加者の印象の向上が課題として残されている。

第5章 複数体エージェントによる第二言語会話支援の検討

本章では、NNS支援エージェント評価実験の結果に基づき、複数体エージェントによる第二言語会話支援手法を検討する。この手法は、2体エージェントの予定された話者交替と修復によって構成される。まず、2体のエージェントが協力してNNSの発話を促すという我々の考えを紹介し、次に、それを実現する手法を説明する。

5.1 提案手法

2.2節で述べた従来研究によると、複数体エージェントによる他人数会話では、参与役割の変化による印象の向上とエージェント間の交替による矛盾の軽減という効果が実験で検証されている。また、対面会話の基本てきな規則と4章で述べた実験結果によると、円滑なヒューマンエージェントインタラクションを達成するために、話者交替の誤りや違反に対処するための修復メカニズムが必要だと考えられる。

第二言語会話支援を目的とする本研究では、それらの知見を利用し、異なる役割をもつ2体の擬人化エージェントからなる第二言語会話支援システムを提案する。2体のエージェントがNS-NNSの二者会話に介入し、発話権のアンバランスの解消を目指す。2体のエージェントをA1, A2とし、次にそれぞれのストラテジー設計を述べる。

5.2 話者交替パターンの設計

5.2.1 A1：発話権譲渡役

NNS支援エージェントの設計と同様である。会話中に他者選択規則も自己選択規則使われなく、自己継続規則も使われない状態が続くと、エージェントが介入する。そしてNNS宛てに次話者選択テクニックを使って発話権を渡そうとする。なお、会話衝突の回避策として、自己開始修復を設置する。具体的に、参加者と発話のオーバーラップが起これば、すぐ発話を中止する。

5.2.2 A2：修復役

A1の介入に対し、介入の次のターンに聞き手（NSとNNS）の反応を判断する。表6の他者開始形式を基準とし、他者開始が起こった度に、A2が会話に介入する。修復の実行を行う上に、NNS宛てに次話者選択テクニックを使う。これによって、トラブル源による発話権譲渡の失敗を回避しながら発話権をNNSに渡すことを期待する。なお、A1と同じく、自己開始修復を設置する。

5.3 発話内容の設計

A1の発話内容は、従来と変わらなく、表1に示した4種類である。

A2の発話内容は、他者修復（前半）と次話者選択（後半）の二つの部分から構成される。

まずは、A1の発話がトラブル源になる場合、いわゆるA1の「NNSの意見を聞く」という発話意図の伝達が失敗する場合、A2が聞き手の立場でその意図を解釈して修復を実行する。

複数体エージェントによる会話システムでは、エージェントの間の話者交替には自己開示（Self-disclosure）が利用されている[18][21]。自己開示とは、自分自身に関連する情報を他者に伝達することである。対人関係における自己開示の機能[35][36]として、感情表出、自己明確化、社会的妥当化が指摘されている。特に社会的妥当化について、多くの場合、自己開示は他者からのフィードバックを促進し、あるいはその他者からの自己開示を受けようとする傾向を強める機能を持つと述べられている。また、鈴木ら[37]は擬人化エージェントからの自己開示によりユーザの自己開示が促されることを実験で検証した。以上に踏まえて、A2が自己開示文を使って修復を実行するように設定した。

次に、A2の次話者選択には、2.1節で述べた話者選択テクニック(a)を使うように設定した。

表9にA2の発話内容を示す。

表9 A2の発話内容

No.	機能	発話内容
1	修復	私も○さんの意見を聞きたいです。
2	次話者選択	どうですか？

表10に、エージェントの介入がトラブル源になる場合の会話例を示す。

表 10 提案エージェントによる会話の一例

A1 : ○さん, どう思いますか? (次話者選択)
NNS : 何? (他者開始)
A2 : 私も○さんの意見を聞きたいです. (他者修復) どうですか? (次話者選択)
NNS : あ, 意見ですか...X と思います.

5.4 プロトタイプと予備実験

NNS 支援エージェントを二体化し, プロトタイプを実装した. 本システムで使用する OS は Windows 10, python 3.6 である. 図 22 に示したように, 2 台の PC に MMD Agent をセッティングした. ソケット通信によって, PC1 にある制御モジュールが 2 体のエージェントの発話を制御できる. 音声検出には, Python 環境で PyAudio[38] というオーディオ I/O ライブラリを利用し, 音声合成には, 音声合成エンジン Open JTalk[39] を利用した.

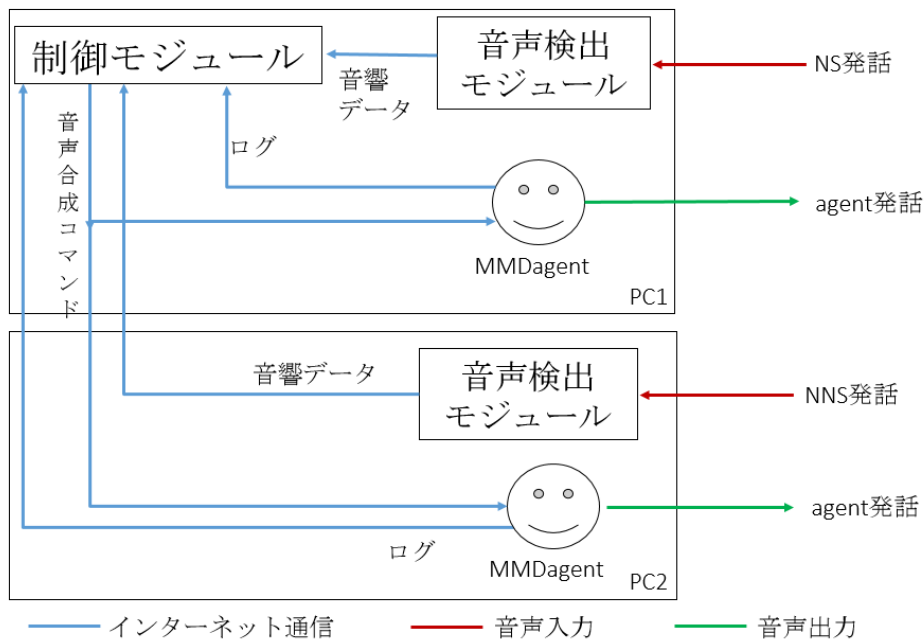


図 22 システム構成

A1 と A2 の自己修復には, MMDAgent の「SYNTH_STOP」コマンドを利用した. モデルの発声を停止させ, 長時間の発話衝突を避けることが可能である. 他者開始の判別には, 選択肢 (1 : 他者開始あり, 2 : 他者開始なし) を設定し, 操作者が表 6 に基づいて決定できるように実装した.

日本語母語話者 1 人と日本語非母語話者 1 人を 1 ペアとし、Woz 法による予備実験を行った（図 23）。会話タスク、実験環境と手順は 3 章と同じようにし、プロトタイプを使って 8 分間の会話を 1 回行ってもらった。



図 23 予備実験の様子

本予備実験は、主観評価の収集を目的とするため、参加者二人を対象としてインタビューを行った。インタビューの質問項目を表 11 に示す。

結果として、質問 1 について、NS が「半分ぐらいはよかった」、NNS が「7 か 8 割ぐらいはよかった。私たちの会話を催促するからいいと思った」と答えた。質問 2 について、NS が「意見を聞きたいという内容はいい」、NNS が「7 割ぐらい理解できた。名前呼ばれたことが分かった。エージェントから意見を求められるのが悪くない」質問 3 について、NS が「2 体の内に 1 体の話が少ないと思った」、NNS が「特になし」と答えた。質問 4 について、NS が「すごく思った。」「前の話を続けるけど、短くして、相手の意見を聞くのが半分だった。つまり話しすぎかなと思って、向こうに話すかなことは半分ぐらい」と答えた。質問 5 について、NNS が「名前が呼ばれたから答えるべきだと思った。」「NS と話してる途中で他の人に向くのがよくないから答えなかった。そしてエージェントの質問に回答しても、私の話への反応が返ってこないからと思った」と答えた。質問 6 について、NNS が「エージェントが自分の話に反応してくれるなら、また質問されるときはそれに答える」と答えた。

表 11 インタビューの質問項目

No.	対象	質問項目
1	NS と NNS	エージェントの介入タイミングについてどう思いますか？
2	NS と NNS	エージェントの発話内容をどの程度理解できましたか？それについてどう思いますか？
3	NS と NNS	2体のエージェントに違いがあると思いますか？（答えが「はい」の場合）何の違いがありますか？
4	NS のみ	エージェントの話を聞いて、相手の意見を聞くように配慮すると思いますか？なぜですか？ （答えが「はい」の場合：エージェントが話した直後に、どれくらい相手の意見を聞きましたか？それはなぜですか？）
5	NNS のみ	エージェントの話を聞いて、それに答えたいと思いますか？なぜですか？（答えが「はい」の場合：エージェントが話した直後に、どれくらい答えましたか？それはなぜですか？）
6	NS と NNS	ほかの意見があれば聞かせて下さい。

今回の予備実験から、参加者が提案システムを用いて会話する際の意識や心理的過程を得たが、個体差による影響が大きい。インタビューの対象人数を増し、多数の参加者に共通の意見を収集できれば、それが提案手法の実装と実験設計の参考になるのではないかと思われる。

第6章 結論

本研究では、会話エージェントによる第二言語会話支援を目的とし、これまでに話者交替規則に基づく NNS 支援エージェントを提案してきた。本稿では、NNS 支援エージェントを用いた評価実験における、エージェントの介入効率や参加者の会話の変化などについて述べた。一定の有効介入ができ、話者交替頻度が高くなったことが分かった。その上に、複数体エージェントによる第二言語会話支援手法を検討した。

謝辞

最後に、本研究を進めるにあたって、始終ご指導ご鞭撻等を頂きました指導教員の井上智雄教授に心より感謝を申し上げます。また、本研究に非常に有意なご意見をくださった副指導教員の三河正彦准教授に深く感謝いたします。

また、二年半の間、最初から最後まで共に頑張って同研究室の皆様、および実験にご協力をくださった参加者の皆様に心から感謝しております。ありがとうございました。

参考文献

- [1]Novinger T. Intercultural communication: A practical guide[M]. University of Texas Press, 2001.
- [2]Zuengler J. Encouraging learners' conversational participation: The effect of content knowledge[J]. *Language learning*, 1993, 43(3): 403-432.
- [3]Zhu W. Interaction and feedback in mixed peer response groups[J]. *Journal of second language writing*, 2001, 10(4): 251-276.
- [4]郭子璇, 唐夢苑, 井上智雄. 非母語話者のための母語話者との会話を支援する会話エージェント (ヒューマンコミュニケーション基礎)[J]. 電子情報通信学会技術研究報告= IEICE technical report: 信学技報, 2018, 117(509): 107-112.
- [5]郭子璇, 井上智雄. 発話交替規則に基づく非母語話者支援会話エージェントの利用(ヒューマンコミュニケーション基礎)[J]. 電子情報通信学会技術研究報告= IEICE technical report: 信学技報, 2018, 118(278): 13-18.
- [6]Goffman E. *Forms of talk*[M]. University of Pennsylvania Press, 1981.
- [7]Sacks H, Schegloff E A, Jefferson G. A simplest systematics for the organization of turn taking for conversation[M]//*Studies in the organization of conversational interaction*. 1978: 7-55.
- [8]Schegloff E A, Jefferson G, Sacks H. The preference for self-correction in the organization of repair in conversation[J]. *Language*, 1977, 53(2): 361-382.
- [9]Schegloff E A. Repair after next turn: The last structurally provided defense of intersubjectivity in conversation[J]. *American journal of sociology*, 1992, 97(5): 1295-1345.
- [10]Extempo <http://www.extempo.com/>.
- [11]Kopp S, Gesellensetter L, Krämer N C, et al. A conversational agent as museum guide—design and evaluation of a real-world application[C]//*Intelligent virtual agents*. Springer, Berlin, Heidelberg, 2005: 329-343.
- [12]Anastasiou D, Jokinen K, Wilcock G. Evaluation of WikiTalk—user studies of human-robot interaction[C]//*International Conference on Human-Computer Interaction*. Springer, Berlin, Heidelberg, 2013: 32-42.
- [13]Shaw E, Johnson W L, Ganeshan R. Pedagogical agents on the web[C]//*Proceedings of the third annual conference on Autonomous Agents*. ACM, 1999: 283-290.
- [14]新藤義昭, 松田洋. K-022 3D-CG アニメーションを用いた対話型 e-Learning システムの開発と教材制作技法に関する研究 (教育工学・福祉工学・マルチメディア応用, 一般論文)[J]. *情報科学技術フォーラム講演論文集*, 2009, 8(3): 575-578.
- [15]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[J]. *情報処理学会論文誌*, 2017, 58(5): 967-980.
- [16]Hayashi K, Sakamoto D, Kanda T, et al. Humanoid robots as a passive-social medium—a field

- experiment at a train station[C]//Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on. IEEE, 2007: 137-144.
- [17]Sakamoto D, Hayashi K, Kanda T, et al. Humanoid robots as a broadcasting communication medium in open public spaces[J]. International Journal of Social Robotics, 2009, 1(2): 157-169.
- [18]有本庸浩, 吉川雄一郎, 石黒浩. 音声認識なし対話における対話感の改善のための複数ロボット連携[J]. SIG-SLUD, 2015, 5(02): 76-77.
- [19]高橋達, 神原誠之, 萩田紀博. 高齢者の発話機会の増加を目的としたソーシャルメディア仲介ロボット: 単体・複数台ロボットを利用したソーシャルメディア仲介実験 (クラウドネットワークロボット)[J]. 電子情報通信学会技術研究報告= IEICE technical report: 信学技報, 2013, 113(84): 31-36.
- [20]藤堂祐樹, 西村良太, 山本一公, 等. 複数の対話エージェントを用いた雑談指向の音声対話システム[J]. 電子情報通信学会論文誌 D, 2016, 99(2): 188-200.
- [21]Iio T, Yoshikawa Y, Ishiguro H. Pre-scheduled turn-taking between robots to make conversation coherent[C]//Proceedings of the Fourth International Conference on Human Agent Interaction. ACM, 2016: 19-25.
- [22]Nakano Y, Fukuhara Y. Estimating conversational dominance in multiparty interaction[C]//Proceedings of the 14th ACM international conference on Multimodal interaction. ACM, 2012: 77-84.
- [23]Rubin J. A review of second language listening comprehension research[J]. The modern language journal, 1994, 78(2): 199-221.
- [24]岡本健吾, 吉野孝. 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[J]. FIT2009 情報科学技術フォーラム, 第 3 分冊, 2009: 393-396.
- [25]Liu R, Inoue T. Application of an anthropomorphic dining agent to idea generation[C]//Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. ACM, 2014: 607-612.
- [26]Inoue T. Naturalistic control of conversation by meal: induction of attentive listening attitude through uneven meal distribution in co-dining[C]//Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. ACM, 2014: 601-606.
- [27]高橋徹, 濱崎雅弘, 武田英明. 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[J]. 人工知能学会全国大会論文集, 2001 (0): 80-80.
- [28]ELAN <https://tla.mpi.nl/tools/tla-tools/elan/>.
- [29]Adobe PHOTOSHOP <https://www.adobe.com/jp/products/photoshop.html>
- [30]メイナード, K.泉子. 会話分析[M], くろしお出版社, 1993.

- [31]吉田奈央, 高梨克也, 伝康晴. 対話におけるあいづち表現の認定とその問題点について [J]. 言語処理学会第 15 回年次大会発表論文集, 2009: 430-433.
- [32]Clark H H, Schaefer E F. Contributing to discourse[J]. Cognitive science, 1989, 13(2): 259-294.
- [33]鈴木佳奈. Other-initiated repair in Japanese: Accomplishing mutual understanding in conversation[J]. Unpublished doctoral dissertation, Kobe University, 2010.
- [34]Ameneh Shamekhi, Q. Vera Liao, Dakuo Wang, Rachel K. E. Bellamy, and Thomas Erickson. Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent[C] // Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). ACM, 2018: 391-403.
- [35]DELREG V J. Appropriateness of self-disclosure[J]. Chelune, GJ and Associates, 1979.
- [36]安藤清志. 対人関係における自己開示の機能[J]. 東京女子大学紀要論集, 1986, 36(2): 167-199.
- [37]鈴木聡, 山田誠二. 擬人化エージェントからの自己開示と第三者への自己開示の伝達がユーザに及ぼす影響[J]. 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, 2003, 103(410): 13-18.
- [38]PyAudio <https://people.csail.mit.edu/hubert/pyaudio/docs/>.
- [39]OpenJTalk <http://open-jtalk.sp.nitech.ac.jp/>.