

多人数会話に基づくソーシャルインタフェースにおける 身体配置の調整について

Adjustment of Body positioning in Social Interface based Multiparty Conversation

山際 康貴¹ 上原 孝紀¹ 蔵田 洋平¹ 大島 直樹¹

P. Ravindra De Silva¹ 岡田 美智男¹

Kouki Yamagiwa¹, Takanori Uehara¹, Yohei Kurata¹, Naoki Ohshima¹,

P. Ravindra De Silva¹, and Michio Okada¹

¹豊橋技術科学大学 情報・知能工学系

¹Toyohashi University of Technology Department of Computer Science and Engineering

Abstract: Recently, few researchers have been pay attention to develop robots to participate in and active multipart conversation. Since, in this study our main focused is to explore body positioning of robot when establish the multipart conversation on the table. During the interaction, user can listen to news source (broadcast by three Muu robots) and user can interrupt/interact with Muu. In this consequence, three Muus consider the user's behaviors to change their positioning, conversation, and interactive structure (social display) to coordinate with user. In this paper, we describe the background of the research, mechanism of multipart conversation, and body positioning of Muus to behave as interactive social interface.

1.はじめに

会話では、「話し手」や「聞き手」として直接会話に加わらなくとも、誰かの会話を耳にすることで、情報が得られることがある。例えば、電車の中で「今日は雨が降るらしいから傘が必要だよ」「え、そうなの？今は晴れているのに」という会話を耳にすれば、今日の天気予報についての情報を得たことになる。複数のロボット（エージェント）による多人数会話の場は、この例と同じようなオープンコミュニケーションとしての性質を備えていることが指摘されている。つまり、「傍参与者」としてその会話に直接関わることがなくても、その会話を聞き流しつつ、その会話によって提供される情報の一部を手に入れることができる。もし、その情報に興味があれば「傍参与者」から「話し手」に役割をシフトして会話に割り込むこともできる。

本研究では、このような多人数会話の場をつくり、その場を介してユーザにネットワーク上のニュースに関する情報提供を行うソーシャルインタフェースのプラットフォームとして、Muu(ICD1.01beta)(以下Muuとする)を構築した。

Muu は実世界指向の Web ブラウザとしての応用が期待できる。コンピュータやキーボード・マウス



図1 構築した Muu(ICD 1.01 beta)の外観

などの操作方法に関する知識を必要とせず、私たちが普段行っている会話などのコミュニケーション・スキルを利用するため、会話を通じてシステムを利用することができる。また、一般的な Web からの情報取得の方法は、マウスやキーボードを利用してブラウジングを行うプッシュ型であるのに対し、Muu はテレビ視聴などのオープンコミュニケーション形態におけるプル型の情報取得方法に近いともいえる。このような特徴を活かすことで、子どもや高齢者な

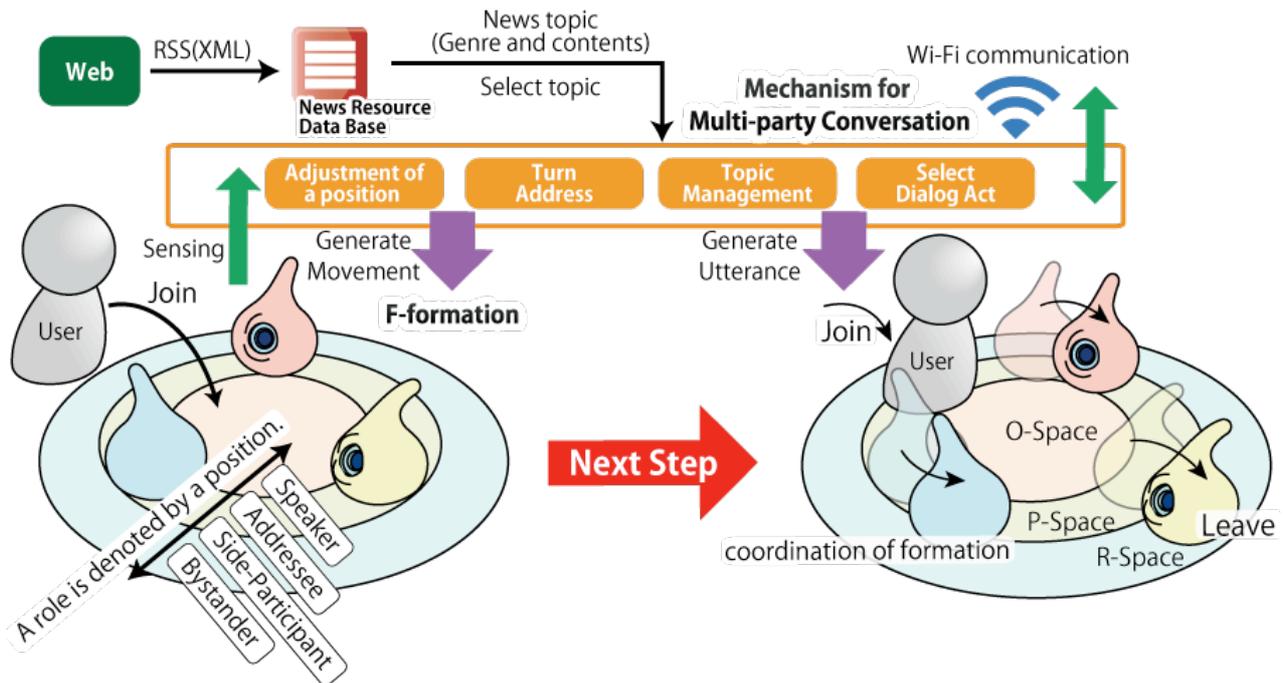


図2 システムの概要

どの情報弱者に向けたユニバーサルインタフェースとしての応用が期待できる[1].

本論文では、構築したプラットフォームのシステム、Muu のハードウェア構成について述べたあと、多人数会話の具体的な生成方法、身体配置の調整方法を述べるとともに、ソーシャルインタフェースとしての可能性について述べる。

2. 研究背景

2.1 多人数会話の場と緩やかな共同性

音声対話システムなどの、人とシステムの1対1の会話では、互いに「話し手」と「聞き手」となることで、会話の場が維持される。発話は相手からの支えを予定して繰り出されるために、一方の応答が得られないと会話が途切れて、会話の場が壊れてしまう。特に1対1の会話では、「会話の場の維持しあう」という共同行為への参加を強いることになり、それぞれの行動を強く制約するものとなる。

一方、システム側を複数のエージェントとした多人数会話では、ユーザが「話し手」あるいは、「聞き手」とならなくとも、エージェント同士が「話し手」、「聞き手」となることで、会話の場が維持される。また、「話し手」や「聞き手」以外に「傍参与者」という参与役割が生まれる。「傍参与者」には、会話参加への自由度が存在し、たとえ会話に不参加であっても会話の場は維持される。

吉池らは、複数のエージェントと人との多人数会話に生まれる「会話参加の自由度」がユーザの心的負荷を軽減させることを指摘しており、このような会話参加の自由度を筆者らは「緩やかな共同性」と呼んでいる[2].

2.2 身体的空間配置調整

Kendonによれば、作業をする個人の身体の前には、「操作領域」という志向空間が存在し、多人数会話においては、この領域同士が重なりあってできた共同空間を、会話参加者が身体的に維持しようと試みるという。この共同空間を維持しようとする身体の空間配置システムはF陣形と呼ばれる[3].

一度形成された陣形に対し、会話参加者が新たに加わったり、すでにいる参加者が会話から離れたりとすると、参加者の身体配置が適宜修正される。例えば、新たな参加者が加わる場合には、共同空間に近寄ってきた参加者が会話の場に入れるよう、新たな会話参加者を加えて円形になるように陣形が調整される[4].

このとき、新たな会話参加者が会話の共同空間に近づくことによって、会話へ参加したいという自らの心的態度の表示を行っている。一方、その表示を受けた既存の参加者たちは、共同空間を新たな参加者のために変化させることで、「あなたを会話に受け入れますよ」という心的態度を新たな参加者に表示しているのである。

会話参加者の身体の配置調整には、心的態度を物理的距離の調整として表示している側面がある。複数エージェントと2人のユーザーによって作られる共同空間では、エージェントからは、ユーザーの身体配置の調整から心的態度を押し量れる可能性がある。同様に、ユーザーがロボットの身体配置の調整を見て、心的態度を押し量れると考えられる。

また、仮想空間上の複数エージェントと人のインタラクションでは、F陣形を考慮することで、エージェントの印象が向上し、ユーザーのストレスが減少することがわかっており、「緩やかな共同性」という観点からもF陣形を考慮したインタラクションについての研究は重要であると言えるだろう。

3. システム概要

構築したシステム全体の簡単な概念図を図2に示す。

本システムはロボット同士の無線通信を中継するサーバ1台と、ロボット3台により構成される。各ロボットは内部にMySQLによるデータベースを持っており、WebからRSS経由で取得したニュース文を保存している。それらのニュース文はセンタリング理論[5]に基づいて選択され、ロボットたちの会話の話題となる。ロボットの発話はDialogActの連鎖として選択され、それに伴って、音声合成による発話生成される。そして、各ロボットはアドレス方向やF陣形における身体配置の調整などを行い、話者交替を行う。この発話交替に伴って、Goffman[6]やClark[7]らによる「多人数会話における参与の枠組み(participation framework)」の中で述べられているような、「話し手(Speaker)」、「聞き手(Addressee)」、「傍参加者(Side-Participant)」などの役割が生まれる。

また、ユーザーがロボットの話題に興味を示し、ロボットたちで構成される場に近づくと、ロボットはユーザーを受け入れるように位置を再調節する。そして、ユーザーが発言すると、ユーザーの発言内容に合わせてロボットたちの話題が遷移していく。

このシステムによって、会話の場を維持する(発話連鎖を組織化する)という基本的な枠組みが提供できる。会話の話題は関連する話題に次々と変わっていき、まるで井戸端会議のように世間話を繰り広げることができる。

3.1. Muuのハードウェア構成

ここでは、構築したMuuのハードウェア構成について述べる。Muuは直径の異なる半球を上下に組み

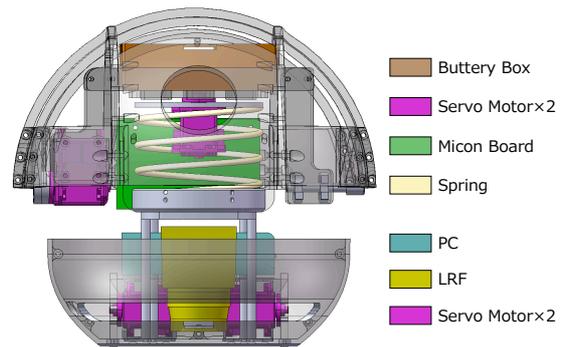


図3 Muuのハードウェア構成

合わせた形をしている。Muuの上半球には志向性を表す大きな目を取り付けられており、Muuがどこを向いているか直感的に分かるようになっている。また、上半球は設置された2つのサーボモータによって、うなずきと首振りを実現できる。一方、下半球には、身体の位置を調整するためのサーボモータと測域センサが取り付けられている。身体の配置を調整するためのサーボモータは、左右に1つずつ設置されており、これにより前進や後退、旋回が可能である。測域センサは、周りのMuuのいる方向や人のいる方向と、その距離を検知できる。なお、Muuにはマイクとスピーカが取り付けられており、簡単な音声認識と発話が行える。

Muuの制御はFitPC2により行われており、ロボットの下部にはFitPC2の放熱のためのファンが取り付けられている。また、Muuたちの会話はWi-Fiを使った無線通信により実現されている。

4. 多人数会話を組織化する方法

本章では、(1)話題の連続性、(2)発話間のつながり、(3)発話交替の秩序、の3つのポイントに着目した多人数会話の組織化手法の概要について述べる。

4.1 話題の連続性

私たちが雑談をするとき、会話参加者の中で話題がある程度共有されつつも、次々と関連話題に遷移していく。雑談においては、「話題の内容を他者に伝える」という側面よりも、話題を移り変えて「会話それ自体を継続する」という側面が強い。

このように、雑談を生成するには話題の連続性やその広がりやを考慮する必要がある。そこで、ここではWeb上で配信されているようなリアルタイムな話題(実世界に密接につながった情報)を利用する。柴田らもWeb上にある文書の利用に着目した雑談

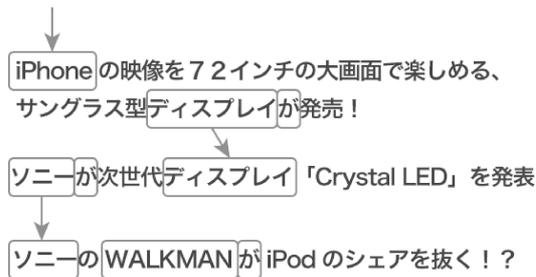


図4 センタリング理論を利用した話題の移り変わり

自由対話の可能性を述べている。

具体的には、RSSで配信されているニュースソースに着目する。RSSとはニュースなどの更新情報を配信するためのフォーマットである。例えば「Yahoo! ニュース—RSS」というページでは、国内外の時事問題や、スポーツ、テクノロジーなどのニュースが1日に数百件ほど配信されている。これらを利用することにより、広がりのある「現実の世界」について話題にすることができる。加えて「時事の話題」という大きなテキスト内の情報であるために、ある程度の連続性が既に確保されている。さらに、RSSはXMLフォーマットから派生したものであるため、情報の取得も容易である。

具体的にどのように話題を移り変えていくか。話題の連続性の確保として、Groszらの提唱するセンタリング理論(Centering Theory)の考えを援用する。センタリング理論では、各発話に話題の中心(center)が1つあると説明される。発話の中の名詞句の中から、話題としての中心性が高いものと低いものに分け、次の話題(名詞句)への移りやすさにランク付けを行う。ランク付けする具体的な手法については、センタリング理論を日本語に適用した事例(下記で示される優先度)に従うものとする。

- (1) ゼロ代名詞
- (2) 「は」が後置する名詞句
- (3) 「が」が後置する名詞句
- (4) 「に」が後置する名詞句
- (5) 「を」が後置する名詞句

例えば、図4における「iPhone」で始まる文章について、この文章には「映像を」というランク5の名詞句と「ディスプレイが」というランクfの名詞句が存在する。上記ランク付けに従い、この名詞句の中でより中心性の高い名詞句(=前向きセンター)であるランク3の名詞句「ディスプレイ」を選ぶ。次の話者に遷移する際、その名詞句「ディスプレイ」を含み、かつ中心性の高い別の名詞句(=後向きセンター)である「ソニーが」を含むような話題を選択することで、話題の連続性が確保できる。

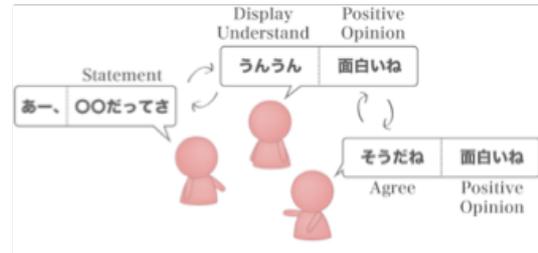


図5 発話間の双方向の機能の例

同様の手順を繰り返すことにより、図4の例における「ソニーの」で始まる文章へと話題を遷移させることができる。

4.2 発話間のつながり

私たちの発話には、伝えたい内容(命題やコンテンツ)とその内容に対する話し手の思惑のようなものが同時に含まれる。例えば「今日の天気知ってる?」と友人に尋ねたとき、「知ってるよ(もしくは知らないよ)」と答えるだろう。このとき、「今日の天気」という内容に対して、「情報の有無を確認したい」あるいは「情報の提供を依頼したい」という話し手の思惑が発話の中に込められている。そして相手は「その問いかけに答える」という行為を遂行する。このように、会話が連鎖する際には、発話自体の内容と違うレベルでの「つながり」が存在している。このような「つながり」を考慮するために、ここでは、談話分析研究におけるアノテーションスキームの一つであるDAMSL(Dialog Act Markup in Several Layers)をその枠組みとして利用する。DAMSLとは、談話分析で用いられる対象の発話に付加された意味をアノテーションする枠組みである。この枠組みによると、発話には2つの機能が含まれている。ひとつは、先行する発話をグラウンドする、BLF(backward-looking function)という機能である。もうひとつは、相手に発話を投げかけるという機能で、FLF(forward-looking function)と呼ばれる。これらの機能は、DA(Dialog Act)と呼ばれており、私たちの発話には話し手に伝えたい内容(話題やニュース・記事の原稿)が含まれるだけでなく、その内容に対する話し手の思惑も同時に伴っていることを改めて強調するような概念と言える。

本研究における発話というのは、この「BLF」と「FLF」のセットであり、発話連鎖というのはこのセットの連鎖したものと考える。例えば図5に示す

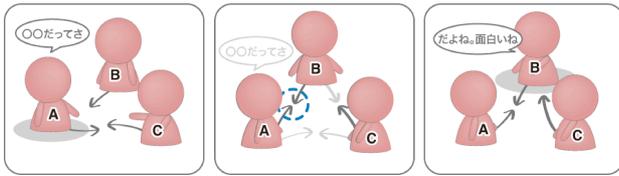


図6 発話交替時のアドレス方向の調整

ように「うんうん、面白いね」という発話は、理解を示すような BLF 機能の DA(Display Understand)と、意見を示すような FLF 機能の DA(Positive Opinion)のセットである。そして、その発話は先行する「OOだってさ」という発話をグラウンドすると同時に、次の発話者の「そうだね、面白いね」という応答を引き出している。

4.3.発話交替のメカニズム

会話の場においては、「一時に1人が、そして1人だけが話す」ことが実現されている。会話参加者の人数に関わらず、誰がどの順番でどのくらいの長さ話すかを前もって決めておかなくとも、順番に1人ずつ話すという秩序がある。会話分析研究において、こうした発話権の移り変わりの秩序は、Sacksらが提唱した「発話交替システム(turn-taking system)」として説明される。そこではターン構成単位(TCU:turn constructional unit)をもとに、発話交替が行われていると考えられている。TCUとは、文、節、句、単語など話し手がターンを構成できる単位のことである[4]。

多人数会話の場においては、TCU末ごとに、次は「誰が話し手になるのか/ならないのか」など、会話参加者間において次話者などの参与役割の選定も必要である。参与役割を決定づけるリソースとしては、視線や身体の向きなどがあると指摘されている。また、榎本らは、3人会話での参与役割の交替において、次のような傾向があることを指摘している。

- 現話者は次話者を見ていることが多く、かつ次話者・非話者は話し手を見ている傾向があること
- 現話者と次話者はターン終了前に視線を交差させる傾向があること

図6はAからBに話し手が交替する際の場面を表したものである。現話者A、次話者B、非話者Cの間でAがCからBにアドレス方向を変更し、(図6でAとBが視線交差し)、Bは次の話し手となる。発話交替のメカニズムの大枠を構成するには、発話単位ごとに誰にアドレスするかという調整がポイントとなる。

ただし、話者交替のタイミングとなるTCUを自動抽出することは困難である。そのため、ニュースソ

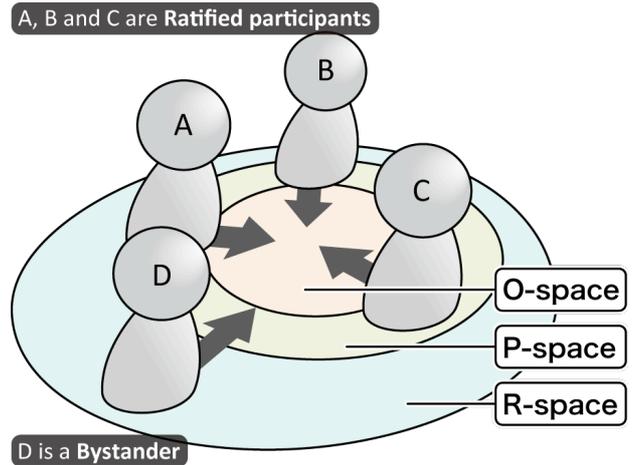


図7 多人数会話におけるF陣形

ースを形態素解析し、文節ごとに分割し、その文節に「ね」などの助詞を加える。こうすることで、ニュースソースを話し言葉で発話しつつ、発話交替が行える。

5.身体配置の調整行動

これまで、筆者らは多人数会話の場を構築する手法について議論してきた。そうして構築した多人数会話の場に、どのようにして人が参与するのか。さらなる「緩やかな共同性」を生み出すために、ここでは、KendonのF陣形概念に基づき、ロボットの身体配置の調整に着目する。

5.1.多人数会話に見られるF陣形

一般に、3人以上での会話に見られるF陣形では、参加者の身体配置は円形になる。この円形は、会話参加者それぞれが身体位置を調整した結果として生まれ、その調整は3つの空間を作り出す(図7)。図では、3人の会話参加者(ratified participant)に傍観者(bystander)が、会話に参加する様子を表している。ここでは、図を用いて、それぞれの空間について説明する。

会話参加者(図中のA,B,C)の中央にできる空間がO空間である。O空間は参加者の「操作領域」が重なってできる空間で、参加者の身体位置の調整によって維持される。そのO空間を維持するために会話参加者が身体を置いている狭い縁がP空間である。その外側の傍観者(図中のD)がいる空間がR空間である。

私たちが、すでに作られた会話の場に入っていく上では、まずR空間で会話を傍観し、時機を見計らってP空間へと身体を運び、参加者へ会話に参加す

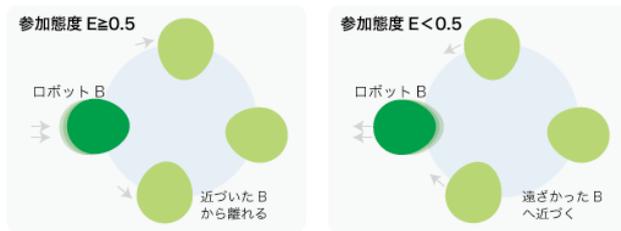


図8 参加態度と身体位置の調整

る意思を表示する。それをほかの参加者が受け入れるのなら、O空間が新たな参加者を考慮した形に調整される。一方、参加者が会話から外れるときには、P空間からR空間へ身体を運び、参加者へ会話に参加しない意思を表示する。参加者がそれを受け入れるのなら、O空間が再度調整される[10]。

このように、身体の調整は会話への参加態度を表すことにもつながる。会話へ参加する者だけでなく、すでに会話の場に加わっている参加者の調整にも注目すると、参加者が、ほかの人の参加や参加者の離脱を受け入れようとしない場合は、O空間は再調整されにくく、受け入れようとする場合には、O空間は再調整されやすくなる。

つまり、ユーザに対して新たな参加や、会話からの離脱を受け入れやすくすれば、「緩やかな共同性」が生まれ、会話参加への自由度が増すと考えられる。

5.2. 身体配置の調整と参加態度

会話の参加者となるMuuは、ロボット下部にある測域センサを用いて、お互いの距離が同じ程度になるように身体位置を調整することで、O空間を維持する。この位置の調整は、ロボットそれぞれが調整するもので、自身の絶対位置を知っている必要はなく、ユーザなどの新たな会話参加者にも対応することができる。

話し手ではないロボットは、O空間を維持しつつも、参加態度の表示として、自身の位置を調整する。興味があるときや、話し手になろうとするときには、前進して会話に参加する態度を示し、興味がないときや、話から離れようとするときには、後退して会話に参加しない態度を示す(図8)。話し手であるロボットは、ほかのロボットのこうした身体位置の変化から、相手の参加態度を計り知ることができる。また、人やロボットが会話の外から新たに参加する場合は、会話に招き入れるように身体位置を調整し、参加を受け入れる(図9)。

このように、参加態度の指標として物理的な距離の調整を取り入れることで、その参加態度はロボットだけでなく、会話に参加するユーザから観察が可

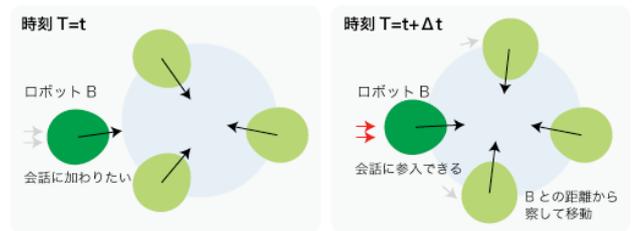


図9 参加者の受け入れにと身体配置の調整

能となる。つまり、心理的態度を周りの参加者に表示しているといえるため、一種の社会的表示といえる。ロボットにしか分からない表示ではなく、ユーザからも観察可能な動作を態度の表示として取り入れることで、ロボットが維持する多人数会話の場がユーザと親和性の高いものとなる可能性がある。

6. まとめ

本論文では、これまで筆者らが「緩やかな共同性」を持った多人数会話に基づくソーシャルインタフェースを実現するための手法についてまとめた。今回は、特に多人数会話に見られる身体配置に注目し、インタフェースに取り入れる方法とその可能性について述べた。今後は、ユーザとの身体配置を考慮したロボットで多人数会話の場を構築し、ソーシャルインタフェースとしての可能性を検証していく予定である。

謝辞

本研究の一部は、科研費補助金(挑戦的萌芽研究24650053)によって行われている。ここに記して感謝の意を表す。

参考文献

- [1] 岡田美智男:『弱いロボット』, ケアをひらくシリーズ, 医学書院 (2012).
- [2] 吉池, 小嶋, De Silva, 岡田:『Mawari:参加メタファに基づくソーシャルインタフェースの提案』, ヒューマンインタフェース学会論文誌, Vol.13, No.1, pp.1-8, (2011)
- [3] Kendon, A.:『Conducting Interaction Patterns of Behavior in Focused Encounters』, Studies in International Sociolinguistics, Vol.7. Cambridge, U.K.: Cambridge University Press, (1990)
- [4] 坊農真弓, 高梨克也共著:『多人数インタラクションの分析手法』, オーム社(2009).
- [5] B. J. Grosz, A. K. Joshi, S. Weinstein:『Centering:a

framework for modeling the local coherence of discourse,
Computational Linguistics』 , Vol. 21, No. 2, pp.203-226
(1995).

- [6] E. Goffman: 『Forms of Talk』 , University of Penn-sylvania Press (1981).
- [7] H. H. Clark: 『Using language』 , Cambridge, Cam-bridge University Press (1996).