

「置いてけぼり」を救う多人数会話活性化ロボット

Facilitation Robot Promoting the Greatest Participation of the Greatest Number in Multiparty Conversation

松山洋一^{1*} 齋藤彰弘¹ 秋葉巖¹ 渡邊萌実¹ 小林哲則¹

Yoichi Matsuyama¹ Akihiro Saito¹ Iwao Akiba¹ Moemi Watanabe¹ and Tetsunori Kobayashi¹

¹ 早稲田大学 基幹理工学部情報理工学科

¹ Department of Computer Science, School of Fundamental Science and Engineering,
Waseda University

Abstract: We propose a conversation facilitation robot aiming at the greatest participation of the greatest number in multiparty conversation with promoting participants motivation. For example, when a participant can not participate in a conversation well, the robot should help him/her starting to speak. In this moment, the facilitation skills should be switched based on situations. In this paper, we propose two skills: 1) *Breaking Barriers* and 2) *Promoting Motivation*. In order to break barrier against participating, the robot should giving the participant chances to get a turn to speak when the participant's barrier is growing higher. In some situations, a participant does not participate because he/she may be bored with the current conversational context. In that situation, the robot is expected to change the current specific topic into wider topics (e.g., "Toy Story" to "animation movies") for them to increase their participation motivation. We conducted an experiment to evaluate the effectiveness of the facilitation skills with 30 participants. Each three subjects were assigned to one group. As the result, speech activities of *Followers* following *Leader* who naturally leads a conversation are significantly increased and silence intervals of the *Followers* are decreased in the proposed system. The result shows that the system gave chances to participate in conversations to the *Followers*.

1 はじめに

多人数の状況において、会話についていけないような「置いてけぼり」になっている会話参加者を救い会話全体を活性化させるようなロボットの開発を行う。ここでいう「会話活性化」とは、会話への参加意欲を向上させ、他者との会話を自己目的的に楽しめるような状況をつくることであると定義する。

近年、多人数会話に会話エージェントを参加させる研究が盛んになってきている。Bohasらは、多人数会話のプロセスをいくつかの層に整理した（Channel層，Signal層，Interaction層，Conversation層）上で，特に低層のChannel層やSignal層に関して，視線や音声情報を用いて現在誰が多人数会話に会話に参加しているかといったEngagementの現象などについて検討した

[Bohus 09]。また最近ではシステムがどのように主体的に発話権を奪取するかといった，ターンテイキングの研究に拡張されている[Bohus 10]。角らは，多種多様なセンサデバイスを用いて，多人数インタラクションのデータを収集しアノテーションを行えるシステムを開発し，多人数インタラクション現象の分析を詳細に行なっている[Sumi 10]。Bohasらの階層モデルに則るならば，これらは主にChannel層，Signal層，Interaction層に注目したものであった。より上位のConversation層に注目した事例としては，堂坂らの会話活性化システムの取り組みが挙げられる[Dohsaka 09]。ここで会話を活性化させるための発話行動として自己中心的な発話と共感的な発話に注目し，それらの発話行動がどのような効果をもたらすか等の検討を行った[Dohsaka 10]。また，Kumarらはテキストを介して会話を行えるキャラクターエージェントを用いた活性化システムの取り組みにおいて，Bales[Bales 50]のSocio-Emotional Interaction Categoriesを用いて発話行動を設計した[Kumar 10]。我々のグループは，過去の実験の分析から質問応答機

*連絡先：早稲田大学 基幹理工学部情報理工学科
知覚情報システム研究室
東京都新宿区早稲田町 27
グリーンコンピューティングシステム研究開発センター
701 室
E-mail: matsuyama@pcl.cs.waseda.ac.jp

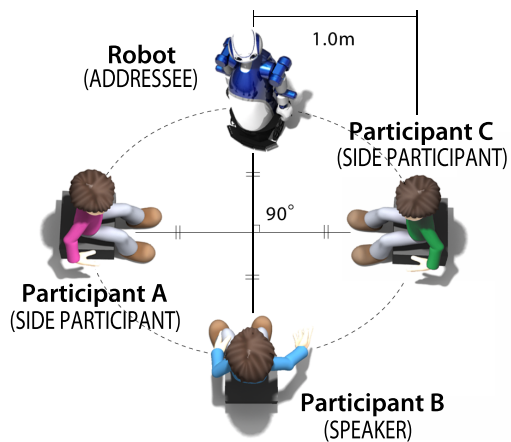


図 1: 本論文での多人数会話の状況．半径 1m の円周上にロボットを含む 4 人の会話参加者．

能に加えて自発的に追加情報を付加するような発話を組み合わせることが活性化に一定の効果があることなどを示してきた [齋藤 11]．

多人数会話には，代表的なタイプとして「会議タイプ」「講義タイプ」「雑談タイプ」のようなものが挙げられる．雑談タイプの会話は，近年音声対話システム研究分野において盛んに研究されている [西村 06]．一般的に雑談は特定の目的を持たないため，他の会話タイプよりも自己目的的な側面が色濃い．本論文では多人数の状況における雑談に注目し，Conversation 層における会話活性化スキルについて検討する．

ロボットを介して活性化させた会話の心理的側面について我々が過去に分析した結果によれば，「楽しさ」「静かさ」「快適さ」などの因子が作用していることがわかった [藤江 11]．客観的な指標としては，活性化させた会話では発話量，笑顔頻度，頷き頻度などが上昇傾向を示すことがわかっている．このうち，本論文では発話量に注目する．会話が活性化している状態を「最大多数の最大発話量」が発せされている状態であるとし，各会話参加者の発話量に注目してロボットの会話活性化効果を評価する．ある会話参加者の発話量が少ない場合でも本人は十分会話を楽しんでいるという状況もあり得るが，「置いてけぼり」を検出し救うという視点においては，発話量は妥当な指標だと考えられる．

以下の章では，2 で会話活性化のために必要なスキルについて検討する．会話活性化スキルを発揮する際に前提となる状況理解として，3 で発話意欲と発話障壁という値を導入する．会話活性化スキルは主にロボットからの自発的質問によって実現されるが，それについて 4 で述べる．5 で会話活性化スキルに必要なトピック制御について述べる．最後に 6 および 7 で会話活性化スキルの効果についての被験者実験の結果と考察を述べる．

2 会話活性化スキル

多人数会話の場を活性化し，「置いてけぼり」となっている会話参加者を救うためのスキルについて検討する．スキルとは，ある目的のために一連の発話行動がまとまった単位と定義する．前述のとおり，会話活性化ロボットの目的は最大多数の最大発話量を引き出すことである．このとき，各会話参加者の発話を妨げる要因について図 1 を例にとりて考えてみる．この図で参加者 C の発話量が極端に少なく，結果的に「置いてけぼり」になっているとする．ここで，参加者 C がどれだけ発話への意欲を持っているかどうかで場合分けできると考えられる．すなわち，参加者 C が現在の話題に関して発話をしたいと思っている場合と，参加者 C がそもそも現在の会話のやりとりへの参加の意欲が低いような場合である．前者の場合，参加者 C に関してさらに大きく分けて 2 つの場合がありうると思われる．1 つ目は，参加者 A と B だけが会話で盛り上がっているときに，参加者 C がその会話に割り込むタイミングを逸しているという場合である．また，誰も発話しておらずしばらくの間無音状態が続いたときに，恥ずかしさなどから話を切り出すのが躊躇されるような場合も考えられる．このような状況において参加者 C を救うためにロボットに期待されるのは，参加者 A と B に配慮しながらも，適切なタイミングで参加者 C に発話の機会を与えることである．参加者 C の会話への参加意欲がそもそも低いような場合は，ロボットは参加者 C の意欲を高めるような行動を試してみるべきだろう．また，会話の場に特に「置いてけぼり」が存在しない場合は，ロボットはいたずらに発話権を取得することは避け，他の参加者の発話の意欲を満たしうるように，傾聴すべきであるだろう．

以上のような考察に基づいて，本研究ではロボットの会話活性化スキルを以下のように分類することにした．

1. *Breaking Barriers* (参加障壁の軽減) : 会話参加者間で著しく発話量の差があり，ある会話参加者が「置いてけぼり」になっている場合は (他の会話参加者の発話の機会を多少奪うことになるとしても) その会話参加者に会話の機会を与えるような発話行動を行う．
2. *Promoting Motivation* (参加意欲の向上) : 「置いてけぼり」となっている会話参加者の会話への参加意欲が低い場合は，その会話参加者に対し意欲を高めるような発話行動を行う．
3. *Attentive Listening* (傾聴) : 会話の場に特に「置いてけぼり」がいない場合は，現発話者の話を傾聴する．

3 参加意欲と参加障壁

3.1 参加意欲

参加意欲とは、現在の話題に関する会話に参加して発話したいという欲求の度合いであり、各会話参加者ごとに定義されるものとする。例えば、ある話題について自分も発話したいという欲求を持っているのであれば、顔の表情の変化や頷きなどの頭部の動きとして現れると考えられる。よって本論文では、会話参加者の顔表情を用いた発話意欲推定手法を検討する。Microsoft Kinect を用いて、1 フレーム 100msec として、顔位置 3 次元 (x 方向, y 方向, z 方向), 顔向き 3 次元 (pitch, yaw, roll), Animation Units 6 次元 (AU0 ~ AU5) の合計 12 次元をフレームごとに取得した。時間的変化を考慮するため、過去 10 フレームの情報を特徴量とし (合計 120 次元), SVM で学習を行うこととした。図 1 のような配置で 4 人着席し (ロボットの位置にも人間が座った), Kinect は各会話参加者の前方 2.5m, 高さ 1.5m の場所に配置した (例えば図 1 の参加者 C の情報を収録する Kinect は参加者 A の背後から C を狙うように配置)。学習データ収集のため映画に関する自由会話を約 10 分ずつ、着席位置を変えながら 1 箇所の座席につき約 30 分ずつ収録した。同時に、会話に参加していない判定者が会話参加者それぞれの参加意欲の有無についてラベル付けを行った。このとき、意欲判定の基準は 1) 視線: 現在の発話者を見ているか否か, 2) 身体位置: 身を乗り出している否か, 3) 頷き: 頷きの頻度が大きい否か, 4) 表情: 笑顔か否か, と定めた。

3.2 参加障壁

参加障壁とは、現在の会話のやりとりへの心理的な参加しにくさの度合いであり、各会話参加者ごとに定義されるものとする。例えば、図 1 の参加者 A と B が活発に会話を行なっている状況がしばらく続いたとき、参加者 C はその会話に参加しにくいと感じている可能性がある。自分以外の会話参加者と自分の発話量の差が大きい状況が長く続いているとき、「置いてけぼり」になっている可能性が高い。よってここでは発話量のみから参加障壁を推定することとした。

発話量は、発話区間が連続している間単調増加し無音区間のときに単調減少する連続的な値とし、自分以外の参加者の発話量の平均と自分の発話量の差を参加障壁とする。また、全員が無音区間が続いた場合もそれぞれの参加障壁が増加するようにする。

4 自発的発話

会話を効果的に活性化させるためには状況に応じてロボット自ら発話行動をしなければならない。このとき「いいタイミング」で「いい質問」をすることがより効果的であると考えられる。ここでは 2 で分類した会話活性化スキルに応じて、ロボットが自らのタイミングで自発的発話を行う仕組みを検討する。

3 で定義した参加障壁が一定以上の高い値を示すとき、その会話参加者が「置いてけぼり」になっていると判断し、その人を質問対象者とする。次にその質問対象者の参加意欲に注目する。参加意欲が高い場合は、その人は現在の話題について興味を持っていると考えられる。質問対象者は現在の話題に興味を持っているのであるから、ロボットはその話題に対して回答しやすい質問をするべきだろう。これは *Breaking Barriers* (参加障壁の軽減) スキルに該当する。例えば現在の話題が「ローマの休日」であるならば、「C (質問対象者)さんは、ローマの休日観たことあるの?」などのような Yes/No 型の質問を行うことにする。

一方、質問対象者が決定された時点で参加意欲が低い場合は、その人は現在の話題について特に興味を持っていないと考えられる。このとき、ロボットは他の会話参加者にも配慮しつつ、少し話題の抽象度を上げることで質問対象者の参加意欲を向上させることを試みる。これは *Promoting Motivation* (参加意欲の向上) スキルに該当する。例えば、質問対象者が「ローマの休日」という話題には特に興味を示していない場合 (しかし他の参加者は興味を持っている)、「C (質問対象者)さんは、恋愛映画は好き?」のように、話題の抽象度を少し上げることで質問対象者の興味を引ける可能性が高くなると考えられる。この話題選択の方法については 5 で述べる。

5 トピック制御

5.1 文書収集

4 で述べた自発的質問を行うために、会話の文脈をモデル化する。本論文では、文脈をトピック (4 の「話題」に該当) の履歴と定義する。

今回は推定する候補となるトピック数を 100 に設定した。分野は映画に限定し、映画のジャンル (「ハリウッド映画」「アニメーション映画」、等)、映画のタイトル (「トイ・ストーリー」「ローマの休日」、等)、映画監督の名前 (「スティーブン・スピルバーグ」「宮崎駿」、等)、俳優の名前 (「ジョニー・デップ」「オードリー・ヘプバーン」、等) が含まれている。Google 検索でトピック候補語ごとに取得できるサイト群 (今回は上位 64 サイト) を 1 文書として扱い、合計 100 文書収

集する．文書ごとに単語の TF-IDF (Term Frequency - Inverse Document Frequency) を求め，その結果を要素とした単語ベクトルを作成する．

$$TF(w, d) = \frac{C(w, d)}{\sum_w C(w, d)} \quad (1)$$

$$IDF(w, d) = \frac{C(d)}{C_{w \in d}(d)} \quad (2)$$

$$TF-IDF(w, d) = TF(w, d) \cdot IDF(w, d) \quad (3)$$

ここで， $C(w, d)$ は文書 d 中の単語 w の出現数， $C(d)$ は文書総数， $C_{w \in d}(d)$ は単語 w の出現数が 1 以上の文書数である．単語 w_i の各文書 d_j での TF-IDF 値 a_{ij} から，単語ベクトル w_i を要素とした文書単語行列 A を作成する．単語間の類似度を単語ベクトル間距離で求める．

$$A = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad (4)$$

5.2 トピック間の類似度計算

100 個のトピック候補語同士の関連度を算出する前処理としてノイズ (文書中の重要でない単語) の影響を抑えるために主成分分析を行い，行列 A を以下のように次元圧縮を行う．

$$A = U \Sigma V^T \quad (5)$$

ここで， U は $m \times r$ 行列， V^T は $r \times n$ 行列である． Σ は $r \times r$ の対角行列であり，固有値 σ_i ($\sigma_1 > \cdots > \sigma_r > 0$) が要素となっている．ここで以下のような関係が成り立つ．固有値の低い部分は削減し，それぞれ $r \times k$ 行列， $k \times k$ 行列の $U_k \Sigma_k$ と近似できる．

$$AA^T = (U \Sigma)(U \Sigma)^T \quad (6)$$

ここで $w_i = [b_1, \dots, b_k]$ を利用して単語同士の関連度を以下のコサイン類似度で求める．

$$\cos(w_i \cdot w_j) = \frac{w_i \cdot w_j}{|w_i||w_j|} \quad (7)$$

5.3 上位トピックの選択

例えば，現在のトピックが「天空の城ラピュタ」であるとする．「置いてけぼり」の会話参加者がそのトピックに興味をなさそうに見える場合，ロボットは他の会話参加者にも配慮しつつも「置いてけぼり」の会話参加者が興味を持てそうなトピックに変更すべきである (*Breaking Barriers* (参加障壁の軽減) スキルに該当) ．

この状況においては，「天空の城ラピュタ」の上位概念である「スタジオジブリ映画」あるいは「アニメーション映画」などと変更した上で「ジブリ映画は好き？」と質問することで，「置いてけぼり」の会話参加者の参加意欲が向上することが期待できる．すなわち，現在のトピックに関連するトピックの中で「1 階層上のトピック」が選択されることが望ましいのであって，「映画」のような一般度の高すぎるトピックは相応しくない．

このような仕組みを実現するために，まず現在のトピックに関して類似度の高いトピック語を複数候補として選び (今回は上位 10 トピック) ，次に一般性について考慮するために，候補語を DF 値でソートする．続いて DF 値を用いて K-means 法でクラスタリングする (今回はクラスタ数 4 程度) ．ここで現在のトピック「天空の城ラピュタ」が含まれるクラスタよりも 1 段階だけ一般性の高いクラスタから次トピックを選択すれば，理想的には「スタジオジブリ映画」などが選択される．

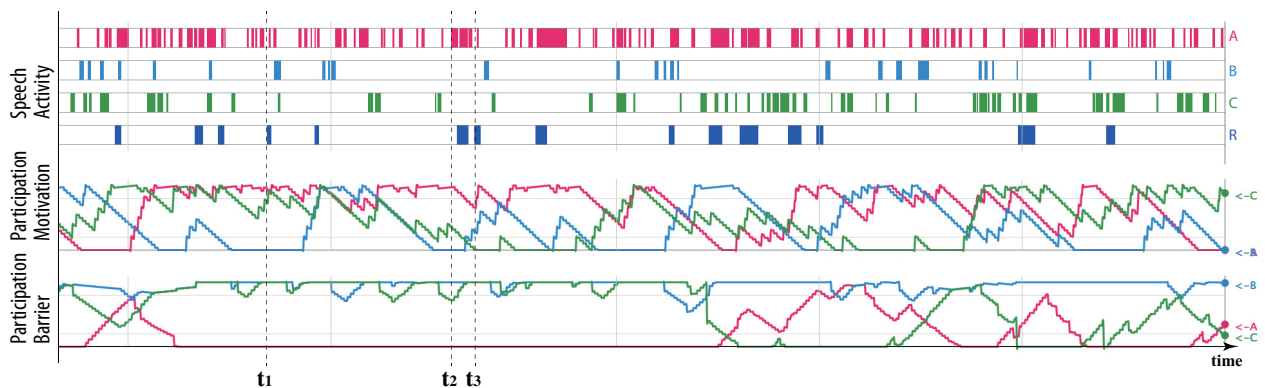
6 実験

6.1 実験条件

会話ロボットの多人数会話活性化スキルの効果を調べるため，以下のような 2 条件を設定した．

条件 1 (baseline): 多人数会話に参加する上で必須の制約を守って行動選択するシステムである．具体的には，トピック推定，参与役割推定，受動的な質問応答が実装されている．参与役割推定については，発話区間情報を用いて現在の発話者 (SPEAKER) ，受け手 (ADDRESSEE) ，傍参与者 (SIDE PARTICIPANT) を推定する．この参与役割の情報を用いて，ロボットの視線を制御する (原則として現在の発話者に視線を向ける) ．質問応答については，現在のトピックを決定した上で，音声認識結果から「5W1H」の質問タイプと「述語 (predicate) 」を抽出し，発話文データベースから検索した結果を返す．例えば，現トピックが「ローマの休日」のとき，ある会話参加者から「どうして好きなの？」という質問を受け取ったとき「どうして (WHY タイプ) 」と「好き (predicate) 」を検索語として「オーディリーが出演しているからね」という発話文を返す．発話文はトピックごとに人手で用意した．本実験では，会話活性化スキルそのものの効果を評価することが目的であるため，トピック推定はオペレータが人手で行うこととした．

条件 2 (proposal): 条件 1 に加えて，2 の会話活性化スキルが加わった提案システムである．参加意欲と参加障壁を推定することで「置いてけぼり」となっている会話参加者を特定し，自発的なタイミングで質問発話を行う．



- t1:** The participant B's participation barrier is high and motivation is low (the participant B seems not interested in the latest topic). The robot actively asks the participant B with a general topic("animation"), "Do you like the animation film?" Then the participant B immediately replies to the robots's question, "Yes, I like animation film."
- t2:** The participant A asks the robot, "Have you ever watched the SAW, SCHEMA?" (The topic is estimated "SAW") Then the robot immediately replies to the participant A with the same topic("SAW"), "I am really lucky to belong this community. I can't live near the murder like him."
- t3:** The participant B's participation barrier is high and motivation is gradually increased (the participant B seems interested in "SAW"). So the robot asks the participant B with the same topic("SAW"), "Have you ever watched the SAW? "

図 2: 条件 2 (提案システム) における実験データの例．上段から各会話参加者の発話量，参加意欲，参加障壁の時系列グラフを表す．

客観的な評価指標として，本実験では発話量と無音区間を調査することとした．本研究の目的は最大多数の最大参加度を促すことであるから，ロボットの発話行動によって会話参加者の発話量が増加し無音区間の平均が減少すること，すなわち発話の機会がすべての会話参加者に与えられることが望ましい．

6.2 実験参加者

被験者は早稲田大学に在籍する日本人の学生 30 名 (男性 16 名，女性 14 名) である．学生の専攻は特定の分野に偏りがないようにし (政治経済，社会科学，教育，文学，医学工学，情報科学等)，年齢層は 19 歳から 39 歳，平均年齢は 22.8 歳だった．

我々の過去の実験の知見から，多人数会話の会話参加者の組み合わせによってロボットの発話行動の効果の現れ方が大きく違うことがわかっている．既知の関係性であったり，初対面でも物怖じしないような会話参加者が多く参加した場合は，ロボットの発話行為が会話の活性化に大きく寄与したり (会話がより盛り上がる)，逆に会話参加者同士で会話が盛り上がりロボットの存在を無視されてしまう状況などが見られることがある．一方，緊張感がある中で会話が始まった場合は，ロボットの発話行為の効果が比較的小さく見逃しやすい (効果はあるが差分が小さい)．よって，ロボットによる会話活性の効果を見るためには，予め会話参

加者の組み合わせを慎重に検討する必要がある．

今回は対人心理の傾向によって 2 つのグループに分けて実験を行うこととした．30 名の被験者は，事前に対人心理尺度に関するアンケートに回答してもらった．アンケートには，菊池 [菊池 88] による KiSS-18 (Kikuchi's Social Skill Scale) を用いた．KiSS-18 は，社会的スキルを身に付けている程度を 18 項目で測定するアンケートである．社会的スキルとは，「対人関係を円滑にはこぶために役立つスキル (技能)」と定義される [堀 01]．KiSS-18 では，6 種類の社会的スキル (初歩的なスキル，高度のスキル，感情処理のスキル，攻撃に代わるスキル，ストレスを処理するスキル，計画のスキル) ごとに質問が用意されている (質問例: 「他人と話していて，あまり会話が途切れない方ですか」，「知らない人でも，すぐに会話が始められますか」，「まわりの人たちが自分とは違った考えをもっている，うまくやっていけますか」等)．それぞれの質問は 5 段階のスケールで回答することが求められる (「5: いつもそうだ」，「4: たいていそうだ」，「3: どちらともいえない」，「2: たいていそうでない」，「1: いつもそうでない」)．今回の平均は 3.27 だったので，3 を基準として社会的スキルが高い傾向を示したグループとそうでないグループに分類することとした．3 名ずつを 1 グループとし，社会的スキルの高いグループを 7 つ (21 名)，そうでないグループを 3 つ (9 名)，合計 10 グループつくった．被験者は互いに初対面である．

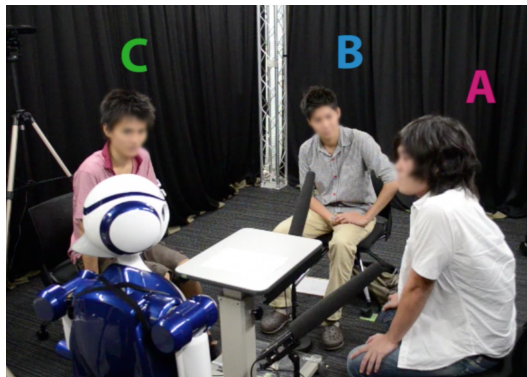


図 3: 実験の様子。映画関連のトピックについて会話参加者同士で会話を行なっている。

6.3 実験手順

被験者は、各実験条件の度に映画関連 50 トピックのリストが書かれた紙を手渡され、そのトピックリスト内から任意のトピックを選んで自由に会話するよう指示された。トピック群は実験条件が変わるごとに差し替えられた（2 条件で合計 100 トピック）。トピックの書かれた紙は実験中でも自由に参照することができた。また、被験者には会話ロボットが会話参加者の一人として参加することが伝えられた。1 グループごとに 2 条件の実施順番を入れ替え、それぞれ 10 分間ずつ実施した。

6.4 実験プラットフォーム

実験プラットフォームとして、会話ロボット SCHEMA（シェーマ）を用いた [Matsuyama 09]。SCHEMA は、身長約 1.2m で、人間が着席したときの目線に合うように設計されている。頭部に 9 自由度（首（ピッチ角，ヨー角），左右の眼球（ピッチ角，ヨー角），左右の眉毛，口），左右の腕部にそれぞれ 6 自由度（肩 2 自由度，肘 2 自由度，手首 2 自由度），合計 21 自由度有している。また、台車により回転・移動することも可能である（本実験では台車の機能を用いていない）。音声認識エンジンは Skood[柴田 02] を用い、音声合成器は本対話システム用に特別にカスタマイズされた株式会社東芝製 ToSpeak を用いた。音声合成時に出力される音素列の情報をもとにリップシンクを行なっている。すべてのモジュールはロボット用ミドルウェア Monea[Nakano 06] によって WiFi を介して統合されている。

6.5 結果

図 3 に実験の様子を示す。

会話参加者の組み合わせに関わらず、すべての会話には会話を主導する人が自然発生的に存在することが観察された。これは会話参加者の社会的スキルの高低に関わらず見られた現象である。本論文ではこのようにグループの会話を主導する人を「リーダー（Leader）」それ以外の人を「フォロワー（Follower）」と呼ぶことにする。図 4 に示すように、会話リーダーは他の会話参加者よりも明らかに発話量が多い。リーダーとフォロワーはそれぞれ複数存在する場合もある。発話量の平均が 3 者で拮抗している場合は、リーダーもフォロワーも存在しないこととする。図 2 は、条件 2 における実験データの例である。この図では、発話量の多い参加者 A がリーダーで B と C がフォロワーである。

発話量: すべての会話参加者の発話量の平均 (sec./min.)

を図 5(a) に示す。ここで、実験実施者のインストラクション不足で意図しない形の会話になった事例（会話参加者間の会話にならずに、ロボットへの集中的な質問応答になってしまった例）やデータが欠損してしまった会話参加者のデータは除いている。以下の結果でも同様の処理を行った。t 検定を実施したところ 2 条件間で優位な差が見られた ($t(26) = 1.77, p < 0.1$)。このとき、ロボットの発話量も条件 1 に比べて条件 2 ではおよそ倍加している（自発的発話の機会が増えているので理にかなっている）。性差の影響は統計的には確認できなかった（男性： $t(13) = 0.74, p = ns$ ，女性： $t(12) = 1.74, p = ns$ ）。図 5(b) は、すべてのグループのリーダーの発話量の平均 (sec./min.) である。2 条件間で優位な差は見られなかった ($t(12) = 0.19, p = ns$)。図 5(c) は、すべてのグループのフォロワーの発話量の平均 (sec./min.) である。2 条件間で優位な差が見られた ($t(13) = 2.28, p < 0.05$)。図 5(d) は、社会的スキルの比較的低いグループのフォロワーの発話量の平均 (sec./min.) である（合計 5 名）。2 条件間で優位な差が見られた ($t(4) = 2.54, p < 0.1$)。

無音区間: 図 6(a) は、すべてのグループの無音区間の平均 (sec./min.) である。2 条件間で優位な差は見られなかった ($t(26) = -1.55, p = ns$)。図 6(b) は、すべてのグループのリーダーの無音区間の平均 (sec./min.) である。2 条件間で優位な差は見られなかった ($t(12) = 1.04, p = ns$)。図 6(c) は、すべてのグループのフォロワーの無音区間の平均 (sec./min.) である。2 条件間で優位な差が見られた ($t(13) = -1.97, p < 0.1$)。図 6(d) は、社会的スキルの比較的低いグループのフォロワーの無音区間の平均 (sec./min.) である（合計 5 名）。2 条件間で優位な差は見られなかった ($t(4) = -0.77, p = ns$)。

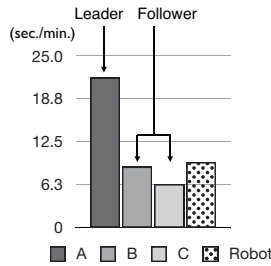


図 4: リーダーとフォロワーの発話量の例

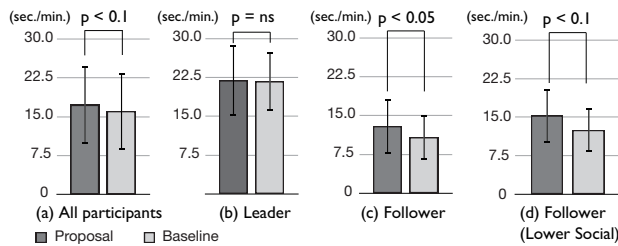


図 5: 発話量の平均 (sec./min.)

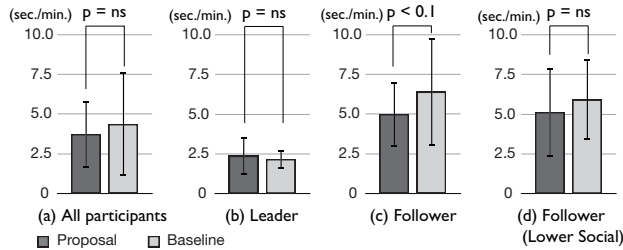


図 6: 無音区間の平均 (sec./min.)

7 考察

リーダーの発話量の平均は条件 1 (baseline) と条件 2 (proposal) であまり変化がないが、フォロワーに関して発話量の平均が上昇、無音区間の平均が減少の傾向を示した。これは、フォロワーに対してロボットの発話行動が発話の機会を与えたといえる。図 2 の時間 t_1 は、Promoting Motivation スキルの様子を示している。参加障壁が高く、参加意欲が低い B に対してロボットがトピックを変更した上で質問を行なっている。図 2 の時間 t_3 は、Break Barrier スキルの様子を示している。この時点では B の参加意欲が上昇を始めているので、ロボットはトピックを変更せずに現トピックに関する質問をしている。これらのスキルは、一定の効果をあげたといえる。実際、定性的には会話参加者の発話タイミングの分散は大きくなっている（条件 1 においてフォロワーの発話機会の間隔が長かったのに対して、条件 2 では間隔が短くなる傾向にあった）。しか

し現在は、発話量に関してまだ微小な上昇にとどまっている。ロボットの自発的な発話によって、全体の発話機会をバランスよく分散させることができ発話量についても優位な効果は見いだせたが、図 5(c) に示したように平均値の差は大きくない（1 分間あたり数秒の差）。これは、ロボットの発話直後にフォロワーの発話量は瞬時的に上昇するものの、それを長く継続させることはできなかったためである。これには、2 つの理由が考えられる。1 つ目は、質問の後の応答である。フォロワーの発話量を大幅に増やすためには、一度フォロワーに質問したあとのフォロワーの応答に対する対応が必要である。例えば、フォロワーに対して「アニメーションは好き？」とロボットが問いかけた場合、当然フォロワーからそれに対する応答が帰ってくると期待されるが、その応答を踏まえた上でロボットが質問を投げかけたり、情報を提示する等の対応をすることによって、フォロワーの参加障壁を下げて参加意欲を上昇させ、発話がより活発なものになると考えられる。また、意欲の高いひとにはより多く語ってもらえるように、質問のタイプを状況に適應すべきである。例えば、意欲の高いひとには Yes/No 質問だけではなく、「その映画のどういうところが好きなの？」といったような、より内発的な発話を促すために質的な質問をすべきだろう。2 つ目は、会話参加者の質問が理解できないときの対応である。例えば、ロボットに対して質問がされたときに即座に反応がないと、会話参加者のロボットへの信頼性 (believability) が揺らぎ、以後の発話を諦めてしまう。これについては、音声認識誤りやコンテンツの不足などによって回答できない場合の内部状態の表出、あるいは不明点の確認発話について検討すべきである。

本システムが潜在的にもっている可能性の一つとして、社会的スキルの比較的低いグループにおける「置いてけぼり」状態になっている人を救うことがあげられる。そのようなグループにおいてこそ、ロボットは積極的な振る舞いが期待される。今回の実験では、社会的スキルに関しては高いグループと低いグループの大きな違いを見いだせなかった。これは被験者数が少ないことも原因のひとつである。今後、社会的スキルの比較的低いグループにターゲットを絞り、被験者数を増やした上で再度実験を行う予定である。また、社会的スキルの高いグループの中に比較的低いフォロワーが一人参加したときに、ロボットがフォロワーを救えるかについても評価する予定である。

8 むすび

本論文では、多人数会話において「置いてけぼり」となっている人を救う会話活性化ロボットの開発を行っ

た．ロボットの目的を最大多数の最大発話量を目指すこととし，会話の状況に応じて 1) *Breaking Barriers* (参加障壁の軽減)，2) *Promoting Motivation* (参加意欲の向上)，3) *Attentive Listening* (傾聴) の活性化スキルを切り替えるシステムを提案した．このような活性化スキルの効果について被験者実験を行ったところ，会話グループの中で最も発話量の多いリーダーに効果は見られなかったが，発話量の比較的少ないフォロワーについては効果があると統計的に結論付けられた．

今後の課題としては，活性効果をより大きくするために自発的質問後のやり取りや，質問応答時の不明点確認発話などについて検討を続ける．また，今回は活性化スキルに注目した実験であったためにトピック推定は人手で行ったが，完全に自律的なシステムにするためにトピック推定手法の検討を行う．参与役割推定についても，マルチモーダル情報を用いたより高精度の推定手法を検討する予定である．

9 謝辞

本研究は，日本学術振興会 科学研究費 若手研究 (B) 「複数人会話活性化ロボットの開発と評価」(課題番号 23700239) の補助により行われたものです．また，ロボットの開発にあたり，本対話システム用に特別にカスタマイズされた音声合成器を，株式会社東芝から御提供頂きました．

参考文献

- [Bales 50] Robert F. Bales, *Interaction Process Analysis: A Method for The Study of Small Groups*, Addison-Wesley, Cambridge, MA, 1950.
- [Bohus 10] Dan Bohus, Eric Horvitz, *Facilitating Multiparty Dialog with Gaze, Gesture and Speech*, in *ICMI'10*, Beijing, China, 2010.
- [Bohus 09] Dan Bohus, Eric Horvitz, *Models for Multiparty Engagement in Open-World Dialog*, in *Proceedings of SIGdial'09*, London, UK, 2009.
- [Dohsaka 09] Kohji Dohsaka, Ryota Asai, Ryuichiro Higashinaka, Yasuhiro Minami and Eisaku Maeda, *Effects of Conversational Agents on Human Communication in Thought-Evoking Multi-Party Dialogues*, *SIGDIAL 2009*.
- [Dohsaka 10] Kohji Dohsaka, Atsushi Kanemoto, Ryuichiro Higashinaka, Yasuhiro Minami and Eisaku Maeda, *User-adaptive Coordination of Agent Communicative Behavior in Spoken Dialogue*, *SIGDIAL 2010*.
- [Kumar 10] Rohit Kumar, Jack L. Beuth, Carolyn P. Ros?, *Conversational Strategies that Support Idea Generation Productivity in Groups*, 9th Intl. Conf. on Computer Supported Collaborative Learning, 2011.
- [Matsuyama 09] Yoichi Matsuyama, Hikaru Taniyama, Kosuke Hosoya, Hiroki Tsuboi, Shinya Fujie, and Tetsunori Kobayashi, "SCHEMA: Multi Party Interaction-Oriented Humanoid Robot", *SIGGRAPH Asia 2009 Emerging Technologies*, 2009, Yokohama.
- [Nakano 06] Teppei Nakano, Shinya Fujie and Tetsunori Kobayashi. "MONEA: message-oriented networked-robot architecture." In *Proceedings 2006 IEEE International Conference on Robotics and Automation (ICRA)*, pages 194-199 2006.
- [Sumi 10] Yasuyuki Sumi, Masaharu Yano, and Toyoaki Nishida. 2010. Analysis environment of conversational structure with nonverbal multimodal data. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI '10)*. ACM, NY, USA, pp.44:1-44:4, 2010.
- [菊池 88] 菊池 章夫，*思いやりを科学する—向社会的行動の心理とスキル*，川島書店，1988．
- [齋藤 11] 齋藤彰弘，松山洋一，藤江真也，小林哲則，*会話ロボットの多人数会話活性化戦略とその評価*，*信学技報*，vol. 111, no. 225, SP2011-53, pp. 7-12, Oct. 2011.
- [柴田 02] 柴田大輔，*ワンパストライグラムデコーダにおける単語履歴の束ね処理に関する検討*，*日本音響学会秋期講演論文集*，pp.151-152 (2002)．
- [西村 06] 西村良太，北岡教英，中川聖一，*応答タイミングを考慮した雑談音声対話システム*，*人工知能学会 音声理解と対話処理研究会*，Vol.46, pp.21-26, 2006.
- [藤江 11] 藤江真也，松山洋一，谷山輝，小林哲則，*人同士のコミュニケーションに参加し活性化する会話ロボット*，*電子情報通信学会論文誌 A*，Vol.J95-A No.1, pp37-45, 2011.
- [堀 01] 堀洋道，吉田富二雄，*心理測定尺度集 II*，サイエンス社，2001．