

ヘイパス！会話のキャッチボールを継続させる ボール型ロボットを用いたファシリテーションシステム

フラハティ 陸^{†1} 橋田朋子^{†1}

本研究ではボール型ロボットを用いることで、会話の流れを止めずに、会話のキャッチボールを継続させるシステムを提案する。従来の会話ファシリテーションシステムの多くは、会話に音声で介入するため、会話の流れを乱す可能性がある。さらにこれらは発話頻度や発話量を指標にしているが、一つの事柄を話すために必要な発話量は一定ではないため、発話量の均衡が必ずしも発話機会の公平性に繋がるとは限らない。提案手法はボールの転がる動きで次に喋る人をさりげなく知らせる。さらにどれだけ応答があったかをもとに計算する会話の広がり性の指標を用いることで、公平な会話のキャッチボールを促す。

Hey, Pass! A Facilitating Robotic Ball to Keep the Conversation Rolling

RICHARD J. FLAHERTY^{†1} TOMOKO HASHIDA^{†1}

We propose a system that uses a robotic ball to keep the conversation rolling without interrupting the flow of the conversation. Most existing conversation facilitation systems intervene by voice, which can interrupt the conversation. Furthermore, these systems use frequency and amount of speech as indicators, but since the amount of speech needed to communicate a single thought can vary, balance in the amount of speech does not necessarily lead to fairness in opportunities to speak. The proposed method uses the rolling movement of a robotic ball to subtly show who should speak next. Furthermore, by using “conversational spreadability” as an indicator, which is calculated based on how many responses a person receives, the proposed method facilitates a fair conversation.

1. はじめに

人は会話に夢中になると会話の独占やおいてけぼり状態に自分では気付きづらい [1]。そのため会話を円滑にするファシリテーションシステムの研究が今までに多く行われてきた。最近ではオンラインでのコミュニケーション機会の増加に伴い、円滑な会話を手助けするオンラインツールの [2] [3] も増えており、コミュニケーションの場にテクノロジーが介入することがより広く受け入れられつつある。筆者らはテクノロジーを用いることで会話についての客観的な情報を集め、公平な発話機会を促せるという考えのもと、特に本研究では対面での多人数の会話を想定し、公平な発話機会を促すことを目指す。なお、1人が複数人に発表する場では発話機会の公平性は重要ではないため、会話の場としては会議やブレインストーミングなどを想定する。公平な発話機会を円滑に促すシステムの要件として、会話中に次に喋るべき人を的確に判断することと、会話の流れを乱したり強制したりせずに次に喋るべき人を知らせることが重要だと考える。そこで、周りがどれだけ応答したかをもとに計算する「会話の広がり性」という新しい指標を用いて次に喋るべき人を判断し、ロボットの移動の動作でさりげなく参加者に知らせる手法 (図 1) を提案する。本稿では、提案手法の詳細とシステムの効果を測る実験、主観評価を報告する。

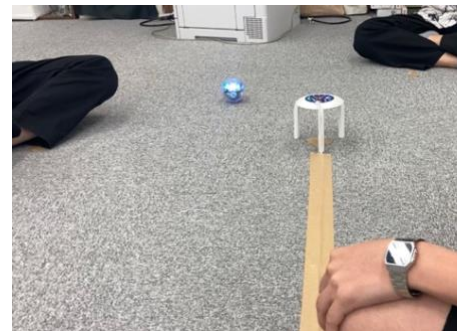


図 1 ヘイパス！のファシリテーションしている様子
Figure 1 Hey, Pass! Facilitating a Conversation

2. 関連事例

本章では本研究に関する事例について、「公平な発話機会を促すツール」と「会話の流れを乱さないための工夫」に分けて紹介する。

2.1 公平な発話機会を促すツールの例

公平な発話機会には会話形式の工夫や各参加者の参加度合いを記録することが必要だと指摘されている [4]。

会話形式を工夫する事例として、アメリカ先住民が議論の際に使っていた発話権を棒として実体化したツールのトーキングスティック [4] [5] がある。棒を持った人のみが喋ることが許可され、喋り終えたら次の人に棒を渡す会話形式で使われる。喋っている際に割り込まれない利点は

^{†1} 早稲田大学
Waseda University

あるが、喋りすぎているかの検知と次に喋るべき人の判断は参加者に任される。本研究は、発話権を実体化する点では同じだが、自律的に移動することで次に話すべき人を的確なタイミングで知らせる点が異なる。

各ユーザーの参加度合いの記録に関係する研究として、共有ディスプレイに発話量を表示することで公平な発話機会を促すことを目指す研究 [1]や、オンライン会議ツールである Zoom 上の会話において、各参加者の相対的発話量を記録し可視化する機能を持つアプリケーションである Macro [2]が挙げられる。さらに従来の多人数の会話ファシリテーションロボット [6]も、発話量をもとに置いてけぼり状態を判断し、適切な手続きを得て発話機会を与える。いずれも発話量を指標にしているが、一つの事柄を話すために必要な発話量は一定ではないため、発話量の均衡が必ずしも発話機会の公平性に繋がるとは限らない。本研究では発話頻度や発話量を指標にせず、発話するとどれだけ周りに会話が広がるかという「会話の広がり性」を指標にする。

2.2 会話の流れを乱さないための工夫の例

従来のファシリテーションロボット [6] [7] [8]の多くは会話に音声で介入し、不自然なタイミングであっても応答を強制するので、会話の流れを乱す可能性がある。

会話の流れを乱さない工夫をしている研究としては、視覚的に情報を提示することで、会話を邪魔せずに会話の特徴を伝える Table Talk Plus [9]や、視線を向かせることで話を振るロボット [10]、呼吸、興味、怯えの動作（ジェスチャー）で会話の特徴を表現する周辺のロボットコンパニオン [11]などが挙げられる。

本研究では、次に喋るべき人をさりげなく知らせる手法として、会話の参加者の近くまでロボットが移動する動作を採用する。音を使わないため、会話の流れを乱さない。物理的なロボットが移動することで、参加者はロボットの側から知らされるだけでなく、手で抑えてもっと喋りたいといった自分の意思を示すことも可能にする。

3. 提案手法

本研究では、会話の流れを止めずに公平な発話機会を促すことが可能な、ボール型ロボットの移動動作を用いたファシリテーションシステム、ヘイパス！を提案する。提案システムでは、公平な発話機会を、どの参加者も同じくらい周りが応答し、応答者に偏りが無い状態と定義する。本章では、このような公平な発話機会を実現するための次に喋るべき人の判断手法と、次に喋る人のさりげない通知に適した移動するロボットの選定についてまとめる。

3.1 次に喋るべき人の判断

次に喋るべき人を的確に判断するために、ページランクアルゴリズム [12]を会話に応用することを考える。ページランクは、検索エンジンで検索結果のページの表示順を

決めるためにサイトの重要度をランキングするアルゴリズムとして考案された。他のサイトからのリンクの合計（入次数）は人気度を表し、他のサイトへのリンクの合計（出次数）は社交性を表す。ページランクにおいては、単に入次数が高いとランクが高くなるわけではなく、リンク元のサイトのランクも考慮する。図2が示すように、リンクが多くて、そのリンク元のサイトの重要度が高いほど重要なサイトだと評価される。サイトに限らず、複数の対象の関係を有向グラフで表すことができれば、ページランクでノードの重要度を計算できる。これはサッカーチームのパスをもとに各選手の重要度を求めるためにも利用される。

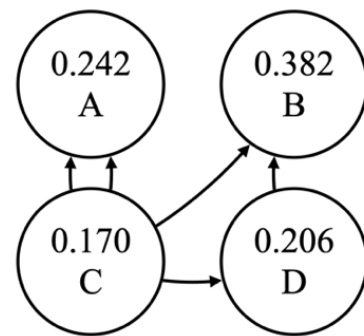


図2 ページランクでランキングされた有向グラフ

Figure 2 A Directional Graph Ranked by PageRank Algorithm

本研究ではサイトのリンクを会話の応答に置き換え、エッジの矢印を応答者から発話者へと設定し、ページランクを用いることで会話の広がり性を計算する。出次数が社交性を示すと似たように、発話者への応答数にさらにページランクを応用したものは会話の広がり性を示す。会話の広がり性は、他の参加者と比べた応答される回数とどれだけ反応される相手に偏りが無いのか、どれだけその先へと会話が続く可能性があるかを表しているため、発話機会を示す指標としての的確と考えられる。

本システムは誰が次に喋るべきかの判断を、会話の広がり性の参加者間の差を抑えるという観点から行う。そのために2つの指標を用いる。一つ目は現発話者に対する応答の回数である。現発話者に対する応答の回数が少ない方に応答してもらうことで応答の偏りを抑える。二つ目に考慮する指標は会話の広がり性である。発話しても会話が終わる参加者を避け、会話の広がり性が高い参加者経由で応答が少ない参加者に話を振る狙いがある。以上から、応答の少なさと会話の広がり性を考慮した下記の式 (1) で算出した値の高い方が次に喋るべき人だと判断する。

$$\left(1 - \frac{\text{現発話者に対する応答数}}{\text{全員の現発話者に対する応答数}}\right) \times \text{会話の広がり性} \quad (1)$$

3.2 転がる動作を用いた次に喋る人のさりげない通知

次に喋るべき人をさりげなく知らせるための空間移動可能なロボットとしては幾つかの形が考えられるが、本研究ではボール型ロボットを用いることとする。ボール型ロボットの移動は転がる動作によって実現され、この動作は会話の輪という狭い空間でも方向転換ができ、急な移動先の変更に対応できるからである。また、ボール型であることで、ロボットの動きを物理的に抑えて自分の番の時間を伸ばすといったユーザの操作も直観的に行うことができる。以上より、提案システムでは現発話者から次に喋るべき人へとボール型ロボットが転がる動作で移動することを繰り返す。

3.3 ヘイパス!の概要

本システムではまず誰が喋っているかを検出し、その人にボール型ロボットが近づく。誰が誰に対して応答したかを検出し、ページランクを用いて会話の広がり性を計算する。現発話者に対する応答の少なさとその会話の広がり性の高さをもとに次に喋るべき人を判断する。現発話者が平均発話時間の半分の時間を喋り続けると、次に喋るべき人にボール型ロボットが転がって移動することでさりげなく参加者に知らせる。次に喋るべき人が喋るとその人が現発話者と認識され、再びサイクルが繰り返される。

4. 実装

ヘイパス!は、音の方向と発話を検出・解析して現発話者と次に喋るべき人を判断する会話解析機構と、ボール型ロボットの転がる動作で次に喋る人を知らせる通知機構からなる仕組みである。

4.1 システム構成

ヘイパス!の具体的なハードウェアとしては、マイクアレイ (ReSpeaker Mic Array v2.0)、制御のためのマイクロコンピュータ (Raspberry Pi 3B+)、ボール型ロボット (Sphero SPRK+) で構成される。マイクロコンピュータの OS は Raspbian Buster, Python のバージョンは 3.7.3, JavaScript のライブラリーは noble v1.7.0, python-shell v3.0.1, sphero v0.9.2 を利用する。

図 3 に示すように、輪になって座っている参加者の中央にマイクアレイを置き、輪の外のマイクロコンピュータと USB で接続する。マイクロコンピュータとボール型ロボットは Bluetooth LE で接続する。初期設定としてまず、起動する際にマイクアレイの下にボール型ロボットを置くことで、座標と向きを合わせる。次に参加者の数と位置を設定する。参加者はマイクから均等な距離でかつお互いに等しい間隔で離れるように、マイクを中心とした円周上に座ってもらう。なお、今回の実装では 3 人での会話を想定する。そのため、各参加者はマイクからおおよそ 70 cm の距離と 60°, 180°, 300° の位置となる。

図 4 に示すように、現発話者と次に喋るべき人を判断する

会話解析機構からの情報をもとに、通知機構が次に喋るべき人を知らせる。以下に各機構についてまとめる。

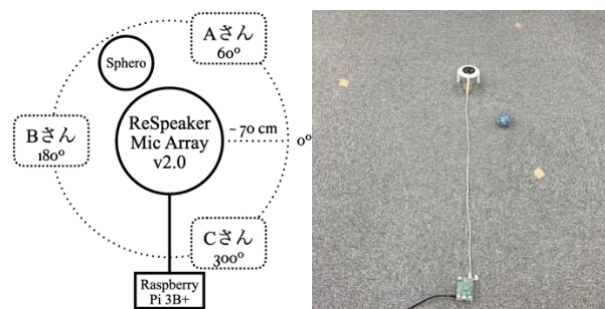


図 3 ヘイパス!の構成

Figure 3 Layout of Hey, Pass!

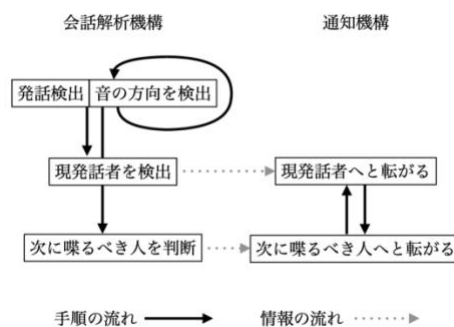


図 4 ヘイパス!の動作の流れ

Figure 4 Operation Flow of Hey, Pass!

4.2 会話解析機構

マイクアレイから取得した音声の情報をマイクロコンピュータで処理することで会話を解析する。図 4 に示すように、マイクアレイからの発話があるかと音の方向の 2 つの情報を常に更新する。発話が検出されたら、参加者が 3 人の場合、0°~119°, 120°~239°, 240°~359°のどの方向から音がしたかで誰が喋ったかを判断する。瞬間的な音による誤認識を省くために 0.3 秒以下の検出された発話は考慮しない。

発話終了後 2 秒間以内に同参加者が再び発話した場合は、連続した発話と判断する。A さんの発話終了後 2 秒間以内に別の参加者 B さんが発話した場合は、B さんが A さんに応答したと見なす。検出された応答で図 5 のような有向グラフを構築する。各エッジの重さは応答の回数を表している。3 章で説明したページランクアルゴリズムをグラフに用いて、各参加者の「会話の広がり性」を計算する。そして 3 章に記載した式で次に喋るべき人を判断する。会話全体を評価するために、会話の広がり性の標準偏差も計算する。

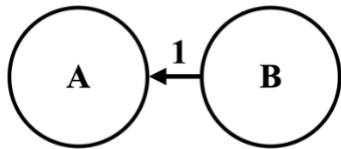


図5 AさんにBさんが応答した際にできるグラフ

Figure 5 The Graph When Person B Responds to Person A

4.3 ボール型ロボットを用いた通知機構

まず、ボール型ロボットは現発話者のもとへと転がっていくことで、発話権を持っていることを知らせる。この状態をリスニングモードと呼ぶ。リスニングモードで現発話者が一定以上の時間を発話し続けたら、ボール型ロボットは次に喋るべきと判断した人の所まで移動することで、そのことを知らせる。この状態をターゲットモードと呼ぶ。次に喋るべき人が1秒以上発話したら再びリスニングモードへと戻る。リスニングモードとターゲットモードを行き来することで、ボール型ロボットの現発話者から次に喋るべき人へと移動する動作を実現する。さらに、現発話者から離れる動作で、そろそろ喋り終えた方がよいことを知らせる。そのため、平均発話時間の半分の時間を喋り続けたら、ターゲットモードに切り替わり、現発話者から離れる。ただし、少なくとも2秒待つように最低時間は設けている。

4.4 動作確認

提案システムを3人の会話において動作させ、正しく発話者を検出し、次に喋るべき人を判断、目的の場所まで移動できていることを確認できた。発話者検出の精度は人の感覚と合っているかで確認した。6分間のシステムを使った会話の動画を再生し、手作業でいつ誰が喋ったかを記録した。その記録とヘイパス!のログを比べ、会話解析機構が手作業で記録した発話の内いくつを検出できたかと、会話解析機構が検出した発話の内いくつが手作業の記録と合っているかを確認した。結果、手作業で記録した発話の内74%を正しく検出でき、発話だと判断した内の89%が手作業の記録と合っていた。一区切り終えた後の沈黙も基本的には正しく検出したが、一度だけ参加者が考え込んで、会話解析機構は連続した発話ではないと判断した。目的の場所にたどり着くかはボール型ロボット内蔵のLEDの色でどこへ向かおうとしているかを可視化することで確認した。発話者の方へと転がる際は白く、各参加者へ転がる場合はそれぞれ違う色に光るように設定した。問題なく目的の場所まで転がることを確認した。

5. 実験

本研究ではどの参加者が発話しても同じだけ会話が広がる可能性がある状態を理想とする。そのため、会話の広がり性の標準偏差が低いことが理想的である。本実験ではボール型ロボットによる通知機構（以下、ボール型ロボッ

トと略す）を使う場合と使わない場合の会話の広がり性の標準偏差を比べる。そして、参加者にはアンケート調査で会話のしやすさと発話機会の公平性を評価してもらう。

5.1 手続き

実験参加者は20代の男女12名である。3人ずつ4つのグループに分けた。3人全員がお互いに喋り慣れている2つのグループと、グループの中に普段喋らない人がいる2グループを設けた。まず、ボール型ロボットなしで6分間会話してもらい、会話の広がり性を算出・記録した。その後6分間、ボール型ロボットを使って会話してもらい同じ会話の広がり性を算出・記録した。事前に、ボール型ロボットは次に喋るべきだと判断した人に転がって行く事を伝えた。

会話のお題は①「色が無い世界または音の無い世界のどちらが良いと思うか」と②「壁の無い家または屋根の無い家のどちらが良いと思うか」に設定した。順番は表1のとおりによりグループによって変え、結果への影響を排除した。

表1 お題の順番

Table 1 Order of Topics

グループ	お題の順番
お互いと喋り慣れているグループ1	①→②
普段喋らない相手がいるグループ1	
お互いと喋り慣れているグループ2	②→①
普段喋らない相手がいるグループ2	

5.2 結果

会話の広がり性の標準偏差をグループごとに図6~9に示す。開始してからのおよそ1分間はまだ会話についての情報が少なく、会話の広がり性の標準偏差が安定しないことから、経過時間1分以降を比較対象とする。

まず、図6~7に示されるように、お互い喋り慣れているグループでは、ボール型ロボットを使った方が会話の広がり性の標準偏差が全体的に低い。一方で図8~9に示されるように、普段喋らない相手がいるグループでは、ボール型ロボットの有無での会話の広がり性の違いは見られない。

以上の結果から、喋り慣れているグループであれば、提案システムによる次に喋るべき人の通知は活かされ、ボール型ロボットは公平な発話機会を促すが、普段喋らない相手がいると、ボール型ロボットからの通知はうまく活かされないことが示唆される。

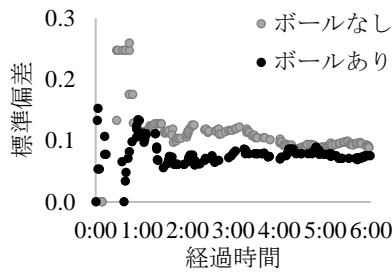


図6 お互いと喋り慣れているグループ1の会話の広がり性の標準偏差

Figure 6 Standard Deviation in Conversational Spreadability from the First Group That Often Talk to Each Other

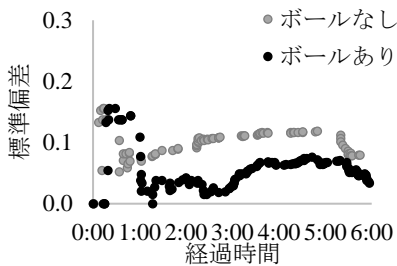


図7 お互いと喋り慣れているグループ2の会話の広がり性の標準偏差

Figure 7 Standard Deviation in Conversational Spreadability from the Second Group That Often Talk to Each Other

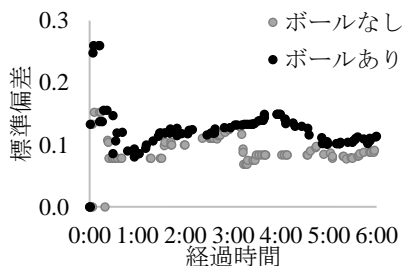


図8 普段喋らない相手がいるグループ1の会話の広がり性の標準偏差

Figure 8 Standard Deviation in Conversational Spreadability from the First Group with a First Timer

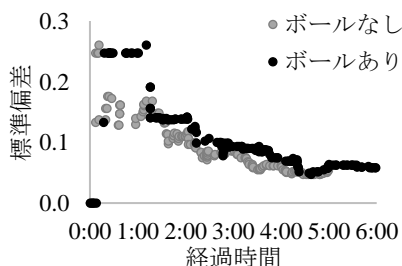


図9 普段喋らない相手がいるグループ2の会話の広がり性の標準偏差

Figure 9 Standard Deviation in Conversational Spreadability from the Second Group with a First Timer

5.3 参加者の評価

ボール型ロボットなしでの会話の後とありでの会話の後の2回、アンケートに回答してもらった。表2の6つの質問項目に5段階で1を「そう思わない」、5を「そう思う」として回答してもらった。結果を分析する際に5と6番の質問項目は、より公平な状態が高い数値と対応するように尺度を逆にした。また自由記述として「システムを活用すると会話に、もしあれば、どのような影響があると思いますか」と「ボール型ロボットはどの程度適切な人へと転がったと思いますか」の2問にも回答してもらった。

表2 アンケートの質問項目

Table 2 Survey Questions

カテゴリー	質問項目
会話のしやすさ	1. 楽に話せた
	2. 会話に集中できた
	3. 会話が弾んだ
発話機会の公平性	4. 全員が同じくらい会話に参加した
	5. 自分あるいは誰か1人が会話を独占した
	6. 自分あるいは誰か1人が置いてけぼりにされた

お互い喋り慣れているグループのアンケート結果を図10に、普段喋らない相手がいるグループの結果を図11に示す。

お互い喋り慣れているグループのアンケートの結果に関して、ボール型ロボットの有無(2水準)と質問項目(6水準)を要因とした2要因の分散分析を行ったところ、質問項目の主効果が有意であり、交互作用は有意ではなかった。今回は被験者数が少なかったこともあり、統計上は有意な差は生じなかったが、図10のグラフからは、「2. 会話に集中できた」「4. 全員が同じくらい会話に参加した」「5. 自分あるいは誰か1人が会話を独占した(逆尺度)」の3項目に関してはボール型ロボットの有無で平均値に違いが生じそうな傾向も見える。今後は被験者増やして検証する必要があると考える。

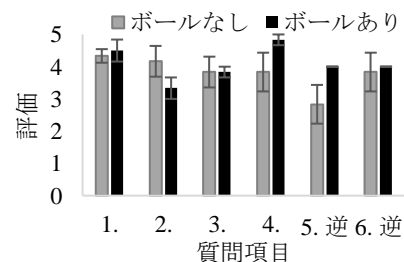


図10 普段喋るグループのアンケート結果

Figure 10 Survey Results of Groups That Often Talk to Each Other

普段喋らない相手がいるグループのアンケート結果に関して、ボール型ロボットの有無（2水準）と質問項目（6水準）を要因として2要因の分散分析を行ったところ、交互作用が有意だった（ $F(5, 25) = 2.87, p < 0.05$ ）。下位検定の結果、「会話に集中できた」の質問項目に関してボール型ロボットの有無の単純主効果が有意であり、平均値からボール型ロボットを使った方が低く評価されたことが分かった。これにより普段喋らない相手がいるグループでは、ボール型ロボットを使うと会話に集中しづらくなることが示唆される。

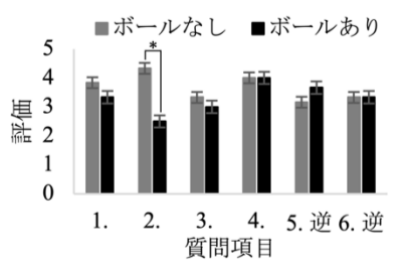


図 11 初対面の参加者がいるグループのアンケート結果

Figure 11 Survey Results of Groups with a First Timer

自由記述の設問のうちシステムの会話への影響について、まずお互い喋り慣れているグループでは、6人中3人からは「会話が弾む」や発話機会が公平になるとの意見があり、3人からは「コントロールされている感じがする」といったコミュニケーションが取りづらいという意見があった。ボール型ロボットの動作の適切性について、6人中4人は適切、1人は大体適切、1人は不安定だと答えた。

一方普段喋らない相手がいるグループでは、システムの会話への影響について、6人中4人からは発話機会が公平になるとの意見があり、2人は「緊張感が生む」「より気遣って会話をする」と答えた。ボール型ロボットの動作の適切性について、6人中3人は適切、3人は大体適切と答えた。

このことからボール型ロボットの動きは大体適切であり、参加者は発話機会を意識させられ、人やグループによってはコミュニケーションが取りづらいと感じるが、発話機会が公平になる可能性があるとし唆される。

6. 今後の展望

本研究では会話の独占と置いてけぼり状態が発生する最低限の人数である3人での実験を行ったが、4人以上の会話ではサブグループに別れるといった状況も想定できる。その場合は、ページランクとは別のグラフ理論の概念である媒介中心性を応用し、サブグループの間を取りまとめる人物を導き出すことが可能だと考える。本研究は自然言語処理を行わないため、どんな言語の会話でも有効に活用できる。非ネイティブ話者はネイティブ話者が多いグループでは置いてけぼり状態になりやすい可能性がある。ヘイパ

ス！はそのような会話で非ネイティブ話者に公平な発話機会を与えられると考える。ただし、言葉を思いつぐのに時間がかかり反応が遅くなるため、現システムの検出方法では応答が正しく検出されない可能性がある。

7. まとめ

本研究では、会話の流れを止めずに公平な発話機会を促すボール型ロボットを用いたファシリテーションシステムを提案した。会話に集中しづらいデメリットはあるが、話を振るのに躊躇がないグループでは、次に喋るべき人を知らせることで公平な発話機会を促せることが分かった。アンケートの自由記述の解答からは、どのグループでもボール型ロボットの通知は適切であったことが分かった。

謝辞 実験のために初対面の相手と会話をしていただいた方々、その他の実験に協力いただいた皆様に謹んで感謝の意を表す。

参考文献

- 1) J. M. DiMicco, A. Pandolfo and W. Bender.: Influencing group participation with a shared display, *Conference on Computer Supported Cooperative Work*, Chicago, (2004).
- 2) Jordan Crook.: Macro just raised \$4.3M to make your never-ending Zoom calls more useful (2020). (引用日: 2022年7月17日). <https://techcrunch.com/2020/07/14/macro-just-raised-4-3m-to-make-your-never-ending-zoom-calls-more-useful/>.
- 3) Katherine Finnell.: Social Collaboration Tools Create Hybrid Work Water Cooler (2022). (引用日: 2022年7月17日). <https://www.techtarget.com/searchunifiedcommunications/feature/Social-collaboration-tools-create-hybrid-work-water-cooler>.
- 4) M. Kornbluh.: Facilitation strategies for conducting focus groups attending to issues of power, *Qualitative Research in Psychology*, (2022).
- 5) K. Fujioka.: The Talking Stick: An American Indian Tradition in the ESL Classroom, *The Internet TESL Journal*, vol. IV, no. 9, (1998).
- 6) 赤川優斗, 福岡維新, 一宮健介, 木村直登, 藤江真也, 小林哲則.: 多人数会話における質疑応答と会話のファシリテーションを行う会話ロボット, *人工知能学会研究会資料*, (2015).
- 7) 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, (2013).
- 8) D. Bohus and E. Horvitz.: Facilitating Multiparty Dialog with Gaze, Gesture, and Speech, *International Workshop on Machine Learning for Multimodal Interaction*, Beijing, (2010).
- 9) 大島直樹, 岡澤航平, 岡田美智男.: Table Talk Plus 多人数会話を媒介するインタラクティブコンテンツとその会話に与える影響について, *電子情報通信学会技術研究報告*, (2009).
- 10) B. Mutlu, T. Shiwa, T. Kanda and H. Ishiguro.: Footing In Human-Robot Conversations: How Robots Might Shape Participant Roles Using Gaze Cues, *International Conference on Human Robot Interaction*, La Jolla, (2009).
- 11) G. Hoffman, O. Zuckerman, G. Hirschberger, M. Luria and T. Shani-Sherman.: Design and Evaluation of a Peripheral Robotic Conversation Companion, *International Conference on Human-Robot Interaction*, Portland, (2015).
- 12) L. Page, S. Brin, R. Motwani and T. Winograd.: The PageRank Citation Ranking: Bringing Order to the Web, *Stanford InfoLab*, Stanford, (1998).