

複数ユーザ対会話エージェントとの 多人数インタラクションの収集と分析

福原 佑貴[†] 中野 有紀子[†]

成蹊大学 理工学部 情報科学科[‡]

1. はじめに

受付や情報案内などの窓口では、複数人の客がグループで案内者と会話することが多い。このようなサービスを実現する会話エージェントシステムを構築するには、複数のユーザを相手に、動的に変化する会話の参与構造 [1] を的確に把握しながら会話を制御する必要がある。

複数のユーザを対象としたコミュニケーションの構造の研究として、[2] では、ユーザの頭部運動と発話の分析から会話構造の確率的推論をおこなっており、[3] では、複数人対話における話者交代の仕組みを視線行動やあいづちと言語情報の関係から分析している。また、情報案内システムの研究では、グループの動的な人数変化を把握して、グループ内のユーザの関係を推定する方法が提案されている [4]。しかし参与構造を把握しながら、状況に応じて複数のユーザに対して適切な情報を提供する会話エージェントの実装を目的とした多人数対話のモデル化はまだ確立されていない。

そこで本研究では、基礎データ収集のために、会話エージェントが 3 人の被験者と会話を行う Wizard-of-Oz による対話収集実験を行い、音声と頭部、および上半身の動作データをモーションキャプチャにより収集した。本稿では、対話収録実験の詳細について述べるとともに、参与構造の推定に向けた基礎的な分析の結果を報告する。

2. 多人数対話データの収集

2.1. 実験の目的

多人数対話の参与構組は、話し手 (Speaker:S)、発話の受け手 (Hearer:H)、会話に参加しているが発話の受け手ではない傍参与者 (Side-Participant:SP)、会話をおこなっているグループ付近にはいるが会話に参加していない傍観者 (Bystander) からなる [1]。本実験ではこれらの参与構造を有する会話データを収集し、会話エージェントとの対話において、複数のユーザ間で動的に変化する対話の役割を分析し、各役割における特徴や差異を明らかにすることである。

2.2. 実験条件

実験条件として、対話への参与の仕方を 2 種類用意し、Wizard-of-Oz エージェントシステムを用いた多人数対話

収録実験を行った。

(A) 役割固定条件：3 人 1 組のグループで週末の外出先を決める会話を行ってもらい、被験者はエージェントに質問することにより場所や施設の情報を得ることができる。3 人のうち 1 人だけ外出できない人を設定することにより、その人の会話参加頻度を下げ、傍参与者となる可能性を高める実験的操作を施した。

(B) 役割変化条件：(A) の状況で会話の途中から外出を希望するユーザがもう一人新規に参加し、途中からこのユーザを加えた 4 人で外出先を決定する。新規参加者がグループ外からグループの会話に新たに加わることに、その役割が傍観者から傍参与者に動的に変化する。尚、実験は各条件 4 回ずつの計 8 回 (各回約 5 分程度) でランダムに実施した。時間内に外出先が決定した場合はその時点で実験を終了とした。

2.3. 収集データ

本実験では 4 人 1 組のグループ、計 3 組 12 人の大学生の参加してもらった。実験のデータを収集するために、会話全体の様子を収録したビデオデータ、ユーザの音声を収録するマイク、身体動作を取得するモーションキャプチャを使用した。また、各課題終了後には会話の中心となった人物の順位付けのアンケートをおこなった。

3. 分析

本実験にて取得したデータのうち今回は実験条件 (A) のデータのみを分析対象とした。データの分析ではビデオデータと音声データから発話区間、発話者、発話の受け手のラベル付けを行った。また、モーションキャプチャから得られた各被験者の頭部と上半身の位置と回転角度の情報により各被験者が誰を注視しているかの推定をおこない、注視対象の自動ラベリングを行った。アノテーションの結果を図 1 に示す。

参与構造と注視との関係を明らかにするために、各発話における各被験者の注視対象を分析した。また、アンケートで得られた会話の中心となった度合いの順位づけをグループ内での優位性順位とし、これと注視との関係についても分析した。

Collecting and Analyzing Multiparty Conversations between
Users and a Conversational Agent

†Yuki FUKUHARA† and Yukiko NAKANO†

‡Seikei University, Faculty of Science and Technology,

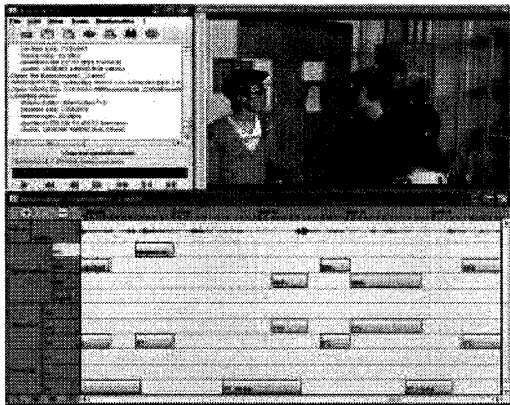


図 1: 発話区間と相手への注視のデータをアノテーションツールにて統合した様子

4. 実験結果と考察

参与構造における役割によって、誰をより多く注視するのかに違いがあるかを分析した結果を表 1 に示す。話し手から受け手 (S→H) が 27%, 受け手から話し手 (H→S) が 24% であり、全体の半分以上を占めている。このことから、話し手-受け手の関係にある参加者は多くの注視を得て、相互に注視を与えあう関係にあるといえる。

一方、傍参与者の注視行動は話し手と受け手に対してほぼ同等であり (SP→S が 13%, SP→H が 14%), これは傍参与者にとって話し手は特別に注目すべき相手ではないことを示唆している。また、S→H が 27% に対して S→SP が 9% であり、話し手は圧倒的に受け手を見ることが多いが (約 3 倍の差)、H→S が 24%, H→SP が 13% であり、その差は 2 倍以下であった。このことより、話し手より受け手の方が傍参与者をより意識していると考えられる。

表 1: 役割と注視対象の関係

H→S	SP→S	S→H	SP→H	S→SP	H→SP
24%	13%	27%	14%	9%	13%

4.2 優位性順位と注視行動の関係

次に、4.1 で得られた参与構造による注視行動の分布が優位性順位によってどのように異なるのかを分析した。各会話データの注視行動を優位性順位ごとに分析した結果を表 2 に示す。優位性順位が 1 位では、話し手として注視行動を行う場合が最も多く (S→H:35%, S→SP: 7%, 計 42%) であり、受け手として注視行動を行う割合も 40% に上る。優位性順位 2 位では、受け手として注視行動を行う場合が最も多く (H→S:25%, H→SP:18%, 計 43%), 話し手として注視行動を行う割合は 37%にとどまる。それに対し、優位性順位が 3 位では、傍参与者としての注視行動の割合が最も高いが (SP→S:16%, SP→H:20%, 計 36%), その他の役割の割合と比較して大きな差はなかつ

た。従って、優位性順位が高いほど話し手として注視行動を行うことが多く、また受け手として会話に参加している割合も高くなっていると考えられる。また、話し手、聞き手としての注視行動の割合が相対的に低い場合は、会話をリードしている可能性は低いと考えられる。

表 2: 優位性順位毎の各役割における注視の割合

	H→S	SP→S	S→H	SP→H	S→SP	H→SP
1 位	28%	9%	35%	9%	7%	12%
2 位	25%	12%	26%	9%	11%	18%
3 位	22%	16%	24%	20%	10%	8%

次に、優位性順位によって、他者から受ける注視量が異なるか否かを調べた結果を表 3 に示す。ここでは、S←H は、その人物が話し手である場合、聞き手から受ける注視の割合を示している。この表から、注視を受ける割合は順位によって大きく変わることはないことがわかる。つまり、受けた注視量の総量だけでは誰が会話の中心的な役割を担っていたのかを予測することは難しく、注視行動から会話の参与構造を予測し、そこから会話をリードしていた人物を割り出す必要があると考えられる。

表 3: 各話者以外の役割の注視の割合

	S←H	S←SP	H←SP	SP←H
1 位	36%	21%	25%	18%
2 位	35%	22%	24%	19%
3 位	37%	23%	23%	18%

5. まとめ

複数のユーザに対して適切な情報を提供する会話エージェントの実現を目指してとして、参与構造と注視行動の関係について基礎的な分析を行った。今後は、参与構造の推定モデルの確立に向けて、さらに詳細な分析を進める予定である。

謝辞: 本研究の一部は文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」(課題番号: 21013042) によるものである。

参考文献

- [1] 坊農真弓: 日本語会話における言語・非言語表現の動的構造に関する研究, ひつじ書房, 2008.
- [2] 大塚和弘: ノンバーバル行動の観測に基づく会話構造の確率的推論, VNV 研究会, 2007.
- [3] 榎本美香: 日本語における聞き手の話者移行適格場の認知メカニズム, ひつじ書房, 2009.
- [4] E. Horvitz and D. Bohus: Open-World Dialog: Challenges, Directions, and a Prototype, JCAI, 2009.