

## D-59 複数人遠隔対話コーパスの構築と取りまとめ役の特徴の分析

波多野翔貴\*, 嶋田和孝\*\*

(\*九州工業大学大学院 情報工学府, \*\*九州工業大学大学院 情報工学研究院)

### 1 はじめに

議論を円滑に行うには取りまとめ役の存在が重要であるが、参加者の誰もがその能力を持つとは限らない。そのため、取りまとめという視点から議論を支援するシステムの研究が進められてきた。既存のシステムは主に対面議論を対象としており、その構築には対面議論コーパスが用いられている。一方で、近年では遠隔でも議論は行われる。遠隔議論では、対話環境の違いにより参加者の行動が対面のものと異なると考えられる。そのため、遠隔議論用システムの構築には遠隔議論コーパスの存在が望ましいが、その数は少ない。

また、対話形式の違いによって生じる取りまとめ役の特徴の差に着目した研究は十分に行われていない。取りまとめ役の言動は議論に大きな影響を与えると考えられるため、対話形式ごとの取りまとめ役の特徴を分析することは重要である。

そこで、本研究では複数人遠隔対話コーパスの構築と、対面と遠隔のそれぞれに対する取りまとめ役の特徴の分析を行う。

### 2 複数人遠隔対話コーパスの構築

Yamamura ら [1] は、全 9 対話の複数人対面対話を対象とした Kyutech コーパスを構築している。この構築手法を基にして、複数人遠隔対話コーパス (Kyutech Remote コーパス) を構築する。なお、本研究で構築・分析するコーパスはプロトタイプ版である。

Zoom<sup>1</sup>を使用し、遠隔議論の実施および収録を行う。4 名 1 組の参加者は、架空のショッピングモールにあるレストラン街において閉店した店の代わりに新規出店させる店を選ぶというテーマで議論を行う。参加者は 20 分を目安に議論を行い、終了後には「誰が議論をコントロールしていたか」などのアンケートに回答する。その後、収録したデータを対象に、著者による対話中の発話の書き起こしを行う。本研究では、計 8 対話分の映像・音声、発話の書き起こしデータを得た。

### 3 取りまとめ役の分類と特徴の分析

Kyutech コーパスと 2 節で構築した Kyutech Remote コーパスを対象として、それぞれの議論形式における取りまとめ役の特徴を分析する。取りまとめ役は、コーパス参加者へのアンケートの回答で、最も多く票を得た話者とする。

まず、著者が発話内容を観察し、取りまとめ役に特徴的な発話について分析を行った。その結果、取りまとめ役は、選択した店の確認といった決定事項の確認を目的とした発話を、議論形式に関わらず行う傾向にあることがわかった。

取りまとめ役の特徴は、発話からの特徴を用いて構築した分類モデルから分析を行えることが示されている [2]。そこで、先行研究の手法を基に、決定木 (CART) を用いて 2 つのコーパスを対象とした取りまとめ役の分類モデルを構築する。各対話の各参加者に対して、コーパスから 5 種類 11 次元の素性を抽出する。使用する素性を次に示す。

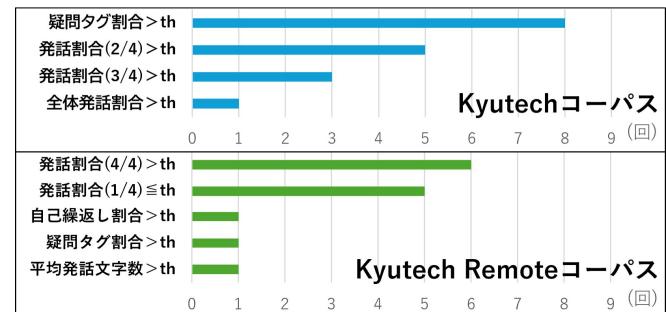


図 1: 対面・遠隔それぞれの取りまとめ役の特徴

- 自身・他者の発話内容を繰り返す発話数の相対的割合<sup>2</sup>
- 内容が疑問を表す発話数の相対的割合
- 対話全体および 4 分割した対話の発話数の相対的割合
- 平均発話文字数、平均発話時間
- 10 秒以上の沈黙後に行われた発話数の相対的割合

Kyutech コーパスに対する 9 分割交差検証と、Kyutech Remote コーパスに対する 8 分割交差検証をそれぞれ行う。構築された分類木による取りまとめ役の分類精度 (F 値) はそれぞれ 0.500, 0.438 であった。

分類実験で構築された全ての木において、取りまとめ役を最も多く含む葉ノードまでの分類条件として使用された素性の出現回数の統計から、取りまとめ役の特徴について分析を行った。図 1 に、使用された素性の統計情報を示す。分類木の各ノードは、分類に用いられた素性 (*th*) の値と、ある閾値との大小関係によってデータを分類する。図 1 から、「疑問タグ割合」と「発話割合」より、対面での取りまとめ役は疑問を含む発話や議論を深掘りする場面での発話が多い可能性が示された。また、「発話割合」より、遠隔での取りまとめ役は話者が議題に対する自分の考えを述べる場面で相対的に発話数が減る傾向にあり、議論を収束させる場面で発話が増える可能性が示された。

### 4 おわりに

本研究では、複数人遠隔対話コーパスの構築と、対面・遠隔それぞれの議論における取りまとめ役の特徴の分析を行った。今後はさらに対話を収録するとともに、コーパスの最終版を完成させ、公開を目指す。

### 参考文献

- [1] T.Yamamura, K.Shimada, and S.Kawahara. The Kyutech corpus and topic segmentation using a combined method. In *Proceedings of ALR12*, pp. 95–104, 2016.
- [2] T.Shiota, T.Yamamura, and K.Shimada. Analysis of facilitators' behaviors in multi-party conversations for constructing a digital facilitator system. In *Proceedings of Collabtech2018*, pp. 145–158. 2018.

<sup>2</sup>対話中で発生した、条件を満たす全発話の数を分母とし、各話者による条件を満たす発話の数を分子とする。

<sup>1</sup><https://www.zoom.com/ja>