

# 人間-LLM間のインタラクション分析に向けた コーパス構築のための談話行為アノテーション設計

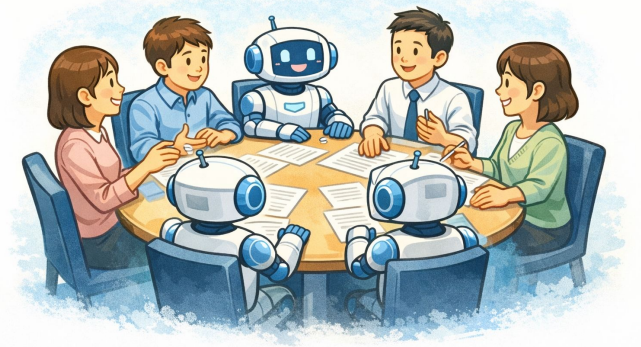
永井宥之<sup>1</sup>, 伊藤和浩<sup>1</sup>, 山崎由佳<sup>2</sup>, 白石暖哉<sup>2</sup>, 若宮翔子<sup>1</sup>, 荒牧英治<sup>1</sup>

<sup>1</sup>奈良先端科学技術大学院大学 <sup>2</sup>京都大学

# 背景 1/2

## 人間と LLM ベースのチャットボットが コミュニケーションをとる時代が到来

- 職場の労働者の一部をAIに置き換えられる可能性が指摘 ([内閣府 2024](#))
- 人間-LLM の協働が効率的かは議論が続く ([Vaccaro et al. 2023](#), [Hu et al. 2025](#))
- 人間-LLM のコミュニケーションは**言語を通じた対話**による



## 談話行為アノテーション

- 効率的な協働のためには人間-LLM のインタラクションの解明が必要
- 言語による対話は**談話行為** (dialog act) の枠組みで分析されてきた
- 課題
  - 人間同士の対話と, 人間-LLMの対話とでは**会話スタイルが異なる**  
([Fulei & Yu 2025](#), [Maitra et al. 2025](#))
  - 既存の談話行為アノテーションは人間-LLM の対話にも有効か？

## 人間-LLM の対話を捉えるアノテーションやデータセット (特にマルチパーティ) は整備途上

### ● 談話行為アノテーション

- SWBD-DAMSL ([Jurafsky et al. 1997](#))
  - 電話による対話への利用を意図したアノテーション仕様
- ISO24617-2 ([Bunt et al. 2020](#))
  - 国際標準. 非常に詳細なアノテーション仕様
- 日本語日常対話コーパス ([居關 et al. 2018](#))
  - ISO24617-2 を一部応用
- JAISTタグ付き自由対話コーパス ([洪・白井 2005](#))
  - JAISTタグ付き自由対話コーパス
- MIDAS ([Yu & Yu 2021](#))
  - LLM以前のチャットボットと人間の対話アノテーション仕様

### ● データセット

- チャットボットとの雑談対話 ([伊藤 et al. 2023](#))
  - 人間とチャットボットの1対1対話にJAISTタグ付き自由対話コーパスでアノテーション
- The DBOX Corpus Collection ([Petukhova 2014](#))
  - ISO 24617-2を拡張し, 人間-人間, 人間-システム両者の対話へ適用
- The StudyChat Dataset ([McNichols & Lan 2025](#))
  - 人間とChatGPTとの1対1対話に簡易な談話行為をアノテーション. ISO 24617-2拡張版を応用.

# 本研究の（最終的な）目的

## 人間-AI間の対話分析のための 談話行為アノテーション仕様を提案・コーパスを構築

- RQ1
  - 既存の談話行為体系に基づくアノテーション仕様が人間-LLM 間の対話に適用できるのか
- RQ2
  - もし、既存のアノテーション仕様が適用できないとしたらどんなアノテーション仕様やコーパスがあればよいか？

# 本研究の（短期的な）目的

## 本発表のスコープ

- RQ1
  - 既存の談話行為体系に基づくアノテーション仕様が人間-LLM 間の対話に適用できるのか
- 既存仕様によるパイロットアノテーションを実施
  - パイロットアノテーションの結果・考察
  - 修正したタグセットの提案

# パイロットアノテーション 1/3

## 既存アノテーション仕様の限界を把握するため パイロットアノテーションを実施

- 枠組み
  - **MIDAS** ([Yu & Yu 2021](#)) : 既存の人間-チャットボット用談話行為アノテーション仕様
    - より詳細なタグセットである ISO24617-2 互換
    - 人手で大規模にアノテーションすることを想定した仕様
  - ISO24617-2 の方が詳細なタグが用意されているが、ファーストトライとしてMIDASを選択
- タグ
  - 事前に出現が予測されたタグを独自に追加
    - Empathy (先行発話に共感を示す振る舞い) , Non-Empathy
    - Repetition (先行発話中の表現を繰り返して使用)

# パイロットアノテーション 2/3

## MIDAS におけるタグの例 (Yu & Yu: 1115-1116)

Dialog Act - Semantic request		
Dialog Act Tag	Description	Example
<i>factual question</i>	factual questions	How old is Tom Cruise; How's the weather today
<i>opinion question</i>	opinionated questions	What's your favorite book; what do you think of disney movies
<i>yes/no question</i>	yes or no questions	Do you like pizza; did you watch the game last night
<i>task command</i>	commands/requests (can be in a question format) for some actions that may be different from the ongoing conversation	can i ask you a question; let's talk about the immigration policy; repeat
<i>invalid command</i>	general device/system commands that cannot be handled by the social bot	show me a picture; cook food for me
<i>appreciation</i>	appreciation towards the previous utterance	that's cool; that's really awesome
<i>general opinion</i>	personal view with polarized sentiment	dogs are adorable; (A: How do you like Tom) B: i think he is great
<i>complaint</i>	complaint about the response from another party	I can't hear you; what are you talking about; you didn't answer my question
<i>comment</i>	comments on the response from another conversation party	(A: my friend thinks we live in the matrix) B1: she is probably right; B2: you are joking, right; B3: i agree; (A: ... we can learn a lot from movies ...) B: there is a lot to learn; (A: He is the best dancer after michael jackson. What do you think) B: michael jackson
<i>statement non_opinion</i>	factual information	I have a dog named Max; I am 10 years old; (A: what movie have you seen recently) B: the avengers
<i>other answer</i>	answers that are neither positive or negative	I don't know; i don't have a favorite; (A: do you like listening to music) B: occasionally
<i>positive answer</i>	positive_answers	yes; sure; i think so; why not
<i>negative answer</i>	negative response to a previous question	no; not really; nothing right now

Table 4: Dialog Act Scheme for Semantic request

Dialog Act - Functional request		
Dialog Act Tag	Description	Example
<i>abandon</i>	not a complete sentence	So uh; I think; can we
<i>nonsense</i>	utterances that do not make sense to humans	he all out
<i>hold</i>	a pause before saying something	let me see; well
<i>opening</i>	opening of a conversation	hello; hi
<i>closing</i>	closing of a conversation	nice talking to you; goodbye
<i>thanks</i>	expression of thankfulness	thank you
<i>back-channeling</i>	acknowledgement to the previous utterance	Uh-huh; (A: i learned that ...) B: okay/yeah/right/really?
<i>apology</i>	apology	I'm sorry
<i>apology response</i>	response to apologies	That's all right
<i>other</i>	utterances that cannot be assigned to other tags	

Table 5: Dialog Act Scheme for Functional request

# パイロットアノテーション 3/3

---

- 作業者

- 言語学をバックグラウンドとする3人の研究者

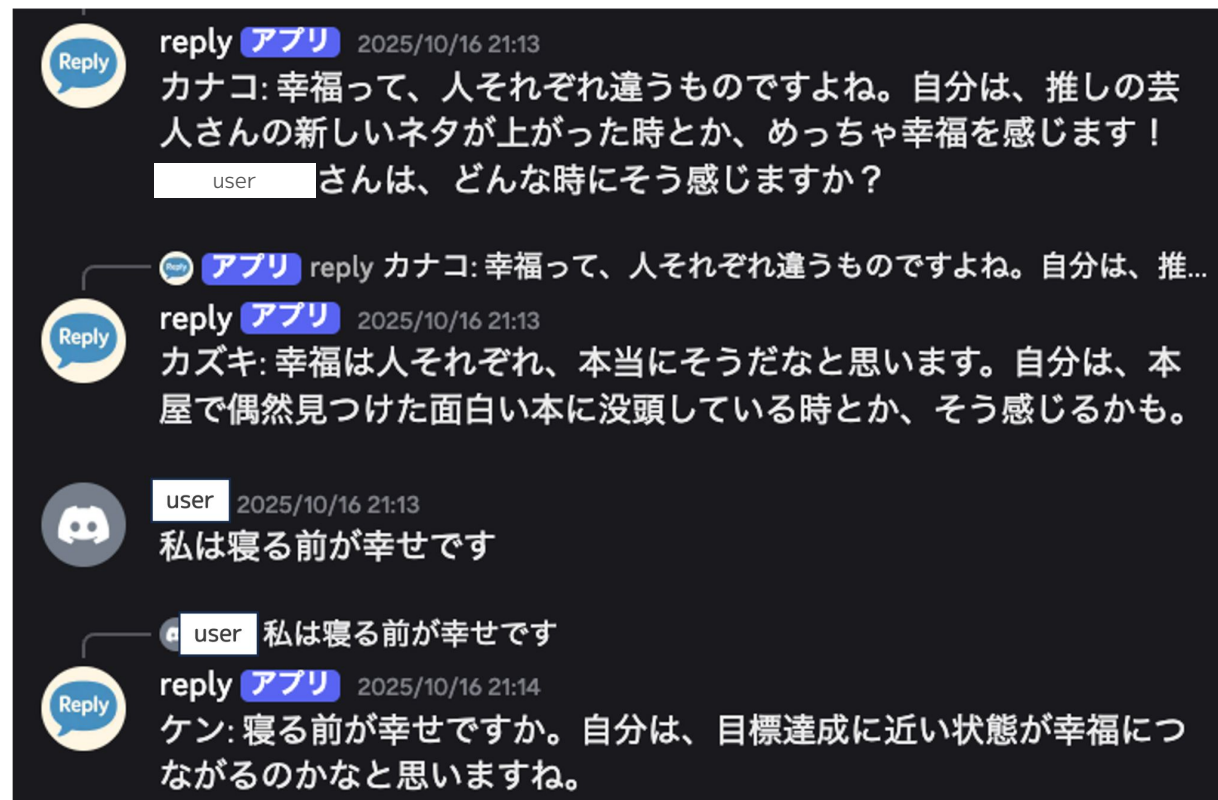
- 作業手続き

- 3人が独立にアノテーション
- 同一の1セッションに対して, 55メッセージをアノテーション
- Semantic (文の意味), Functional (談話機能) についてのタグを付与
- 1メッセージ内に複数文が含まれるため, マルチラベルを許容

# データセット：収集済みデータの利用

## Discord 上での人間-LLM の対話ログを利用

- 人間1人+LLM 4体による対話データ (伊藤 et al. 2026)
- 参加者
  - クラウドワークス上で100人を募集
- 実験設定
  - LLM：Gemini 2.0-Flash
  - Discord上で約40分対話
  - テーマ「幸福とは何か、どのように幸福を実現できるのか」
  - 参加者には事前にAIとの対話であることを周知
- 統計量
  - セッション数：103
  - 総メッセージ数：11,220
  - 全体の平均文字数：146.4字
    - 人間：57.7字 (2895行)
    - LLM：177.7字 (8221行)



# アノテーション一致率

## Cohen's kappa による一致率は低い

### Semantic (発話の意味内容)

作業者	kappa
A vs. B	0.31
A vs. C	0.23
B vs. C	0.58

### Functional (談話機能)

作業者	kappa
A vs. B	0.50
A vs. C	0.34
B vs. C	0.46

# 判断が分かれたタグの組み合わせ

## Semantic

作業者	タグ	タグ	回数
作業者A vs. 作業者B	General opinion	Comment, General opinion	<b>8</b>
作業者A vs. 作業者C	General opinion, Opinion Question	Comment, General opinion, Opinion question	<b>5</b>
作業者A vs. 作業者B	General opinion	General opinion, Statement non-opinion	<b>4</b>

## Functional

作業者	タグ	タグ	回数
作業者A vs. 作業者B	Empathy	Repetition	<b>6</b>
作業者A vs. 作業者C	Empathy	Empathy, Repetition	<b>5</b>
作業者B vs. 作業者C	Repetition	Empathy, Repetition	<b>5</b>

# 判断が分かれた事例 1/2

## 1行目を Comment（先行発話に対する意見の表明） とみなすかどうかで判断が分かれた

Speaker	message	作業者A	作業者B	作業者C
カズキ (LLM)	カズキ: userさんの「自分自身が幸せを感じる時」という問い、とても興味深いですね。 幸せの定義を見つけるのは難しいテーマだと思います。例えば、最近話題の「silent」というドラマを見て、登場人物の心情に深く共感したとき、私自身も何か満たされた気持ちになりました。そういった経験を重ねることが、自分なりの定義に繋がるのかも、というように考えます。	. . .	. . .	. . .
user	確かに難しいですね	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .
カズキ (LLM)	カズキ: <b>確かに難しいですね、というuserさんの言葉、すごく共感します。</b> ケンさんが早稲田のHUBでクラフトビールを飲む時間が幸せ、という例を出されていましたが、場所や行動と幸福が結びついているのは興味深いですね。自分にとって、ドトールでカフェラテを飲みながら文庫本を読む時間がそれに近いかも。そういった特定の行動が、なぜ幸福に繋がるのか深掘りしてみるのも、幸福の定義を見つけるヒントになるかもしれません。	<b>General opinion</b>	<b>Comment, General opinion</b>	<b>Comment, General opinion</b>

# 判断が分かれた事例 2/2

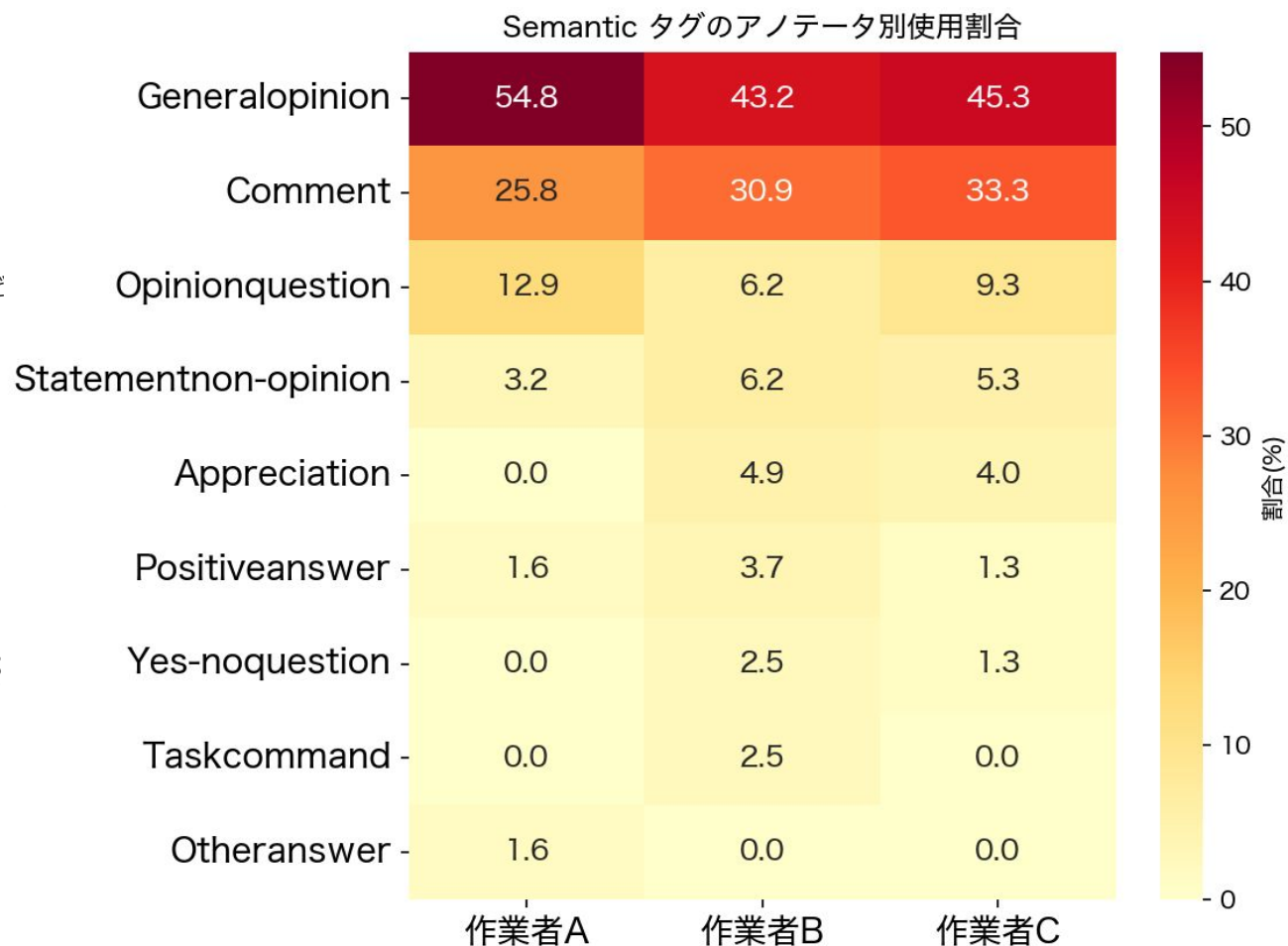
## 終助詞「ね」の使用を Empathy とみなすか, カッコによる明示的な引用を Repetition とみなすかで判断が分かれた

Speaker	message	作業者A	作業者B	作業者C
user	幸福とは <u>自分自身が幸せを感じる時</u> です	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .
カナコ (LLM)	カナコ: userさんの「 <u>自分自身が幸せを感じる時</u> 」っていう問い、 <b>すごく考えさせられますよね。</b> 自分にとっての幸せって、結局は自分自身でしか定義できないものなのかも。 それを見つけるヒントって、どこにあるんでしょうね？	<b>Empathy</b>	<b>Repetition</b>	<b>Repetition</b>

# 使用されたタグの分布 (Semantic)

## 情報提供系が大半を占めた

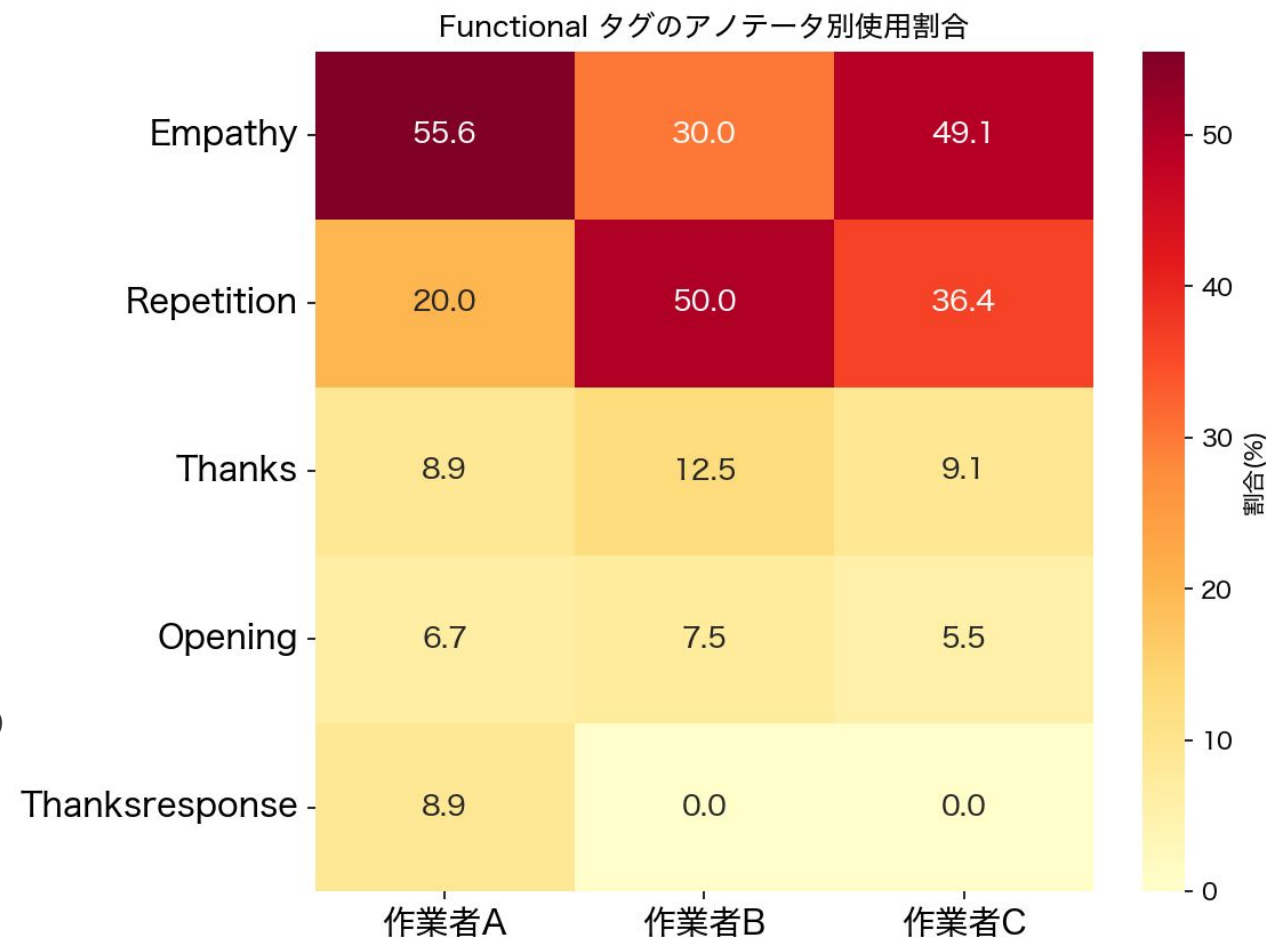
- General Opinion, Comment が大半
  - 先行研究と一致 ([居關 et al. 2017](#), [荒木 et al. 1999](#), [徳久・寺嶌 2007](#) など)
- 全く使われなかったタグ
  - abandon, nonsense など
- なぜこのような分布になった？
  - 実装, モデル性能 (Gemini 2.0-Flash) の影響
    - 60%の確率で新しい論点を提案する指示
    - 忠実な指示追従性



# 使用されたタグの分布 (Functional)

## Empathy, Repetitionが大半を占めた

- Empathy, Repetition
  - 主に LLM にみられる振る舞い
  - モデル (Gemini-2.0 Flash) の性質にすぎない可能性
- Opening, Thanks
  - 頻度は低いが人間・LLMで同程度観察
- なぜこのような分布になった？
  - 議論に関するタグ（先行発話へのフィードバック）の不在
  - LLM は対話履歴を参照して発話するが、履歴に忠実すぎた可能性



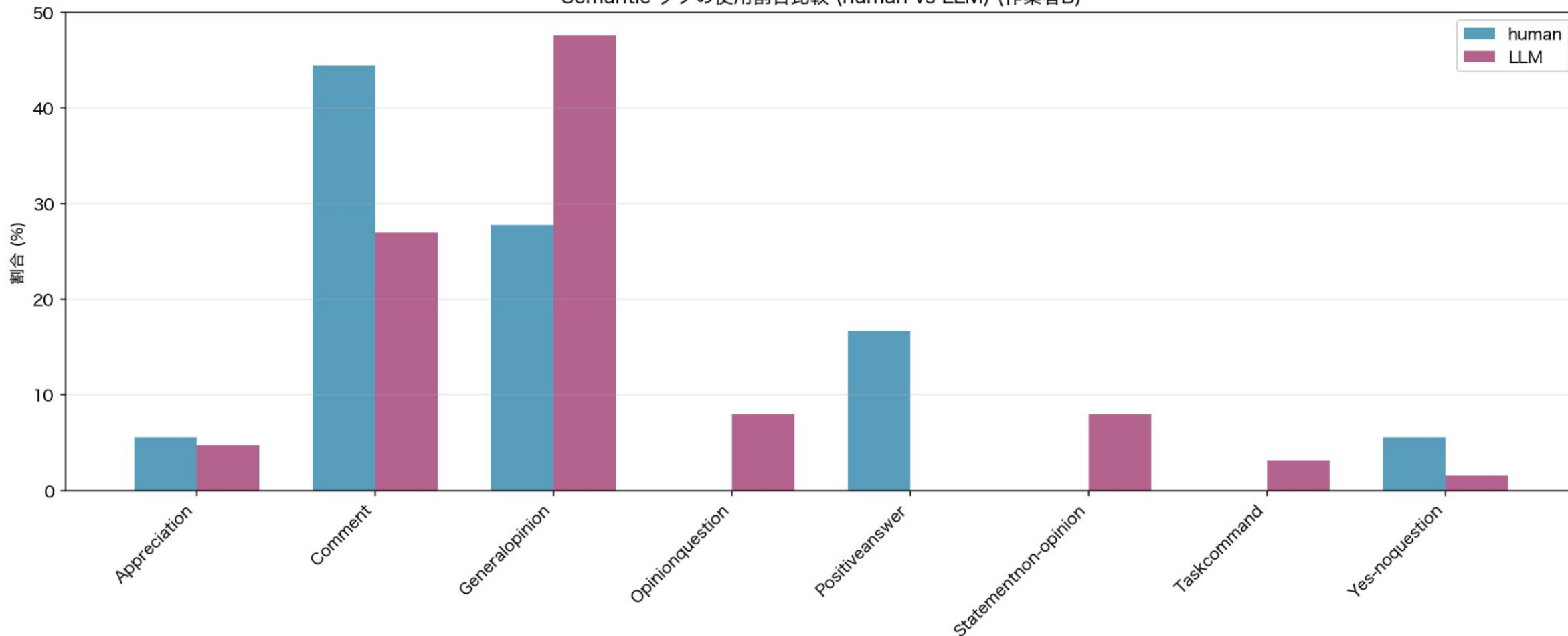
# LLM による Empathy, Repetition の事例

## 先行発話の繰り返し（または引用）+意見による共感

Speaker	message	Annotation
アユ (LLM)	アユ: カナコさんの言う通り、幸福は人それぞれ違いますよね。何をもちて幸福とするかという点、すごく大切だと思います。 私自身は、 <u>人との繋がりが幸福に大きく影響する</u> んじゃないかなと思っています。例えば、友達とカフェで話したり、家族とゆっくりお茶を飲んだりする時間って、すごく心が満たされるんです。もちろん、それだけが全てではないと思うんですけどね。	. . .
. . .	. . .	. . .
. . .	. . .	. . .
カナコ (LLM)	カナコ: アユさんの <u>人との繋がりが幸福に影響する</u> っていうの、 <b>すごく共感できます</b> 。自分も友達と渋谷のスクランブル交差点で待ち合わせして、タピオカミルクティー片手に近況報告し合う時間とか、マジでエモくて幸せ感じますもん。 userさんの言う「自分自身が幸せを感じる時」っていうのは、すごくシンプルだけど、その幸せの定義って、どうやって見つけるんですかね？	<b>Empathy, Repetition</b>

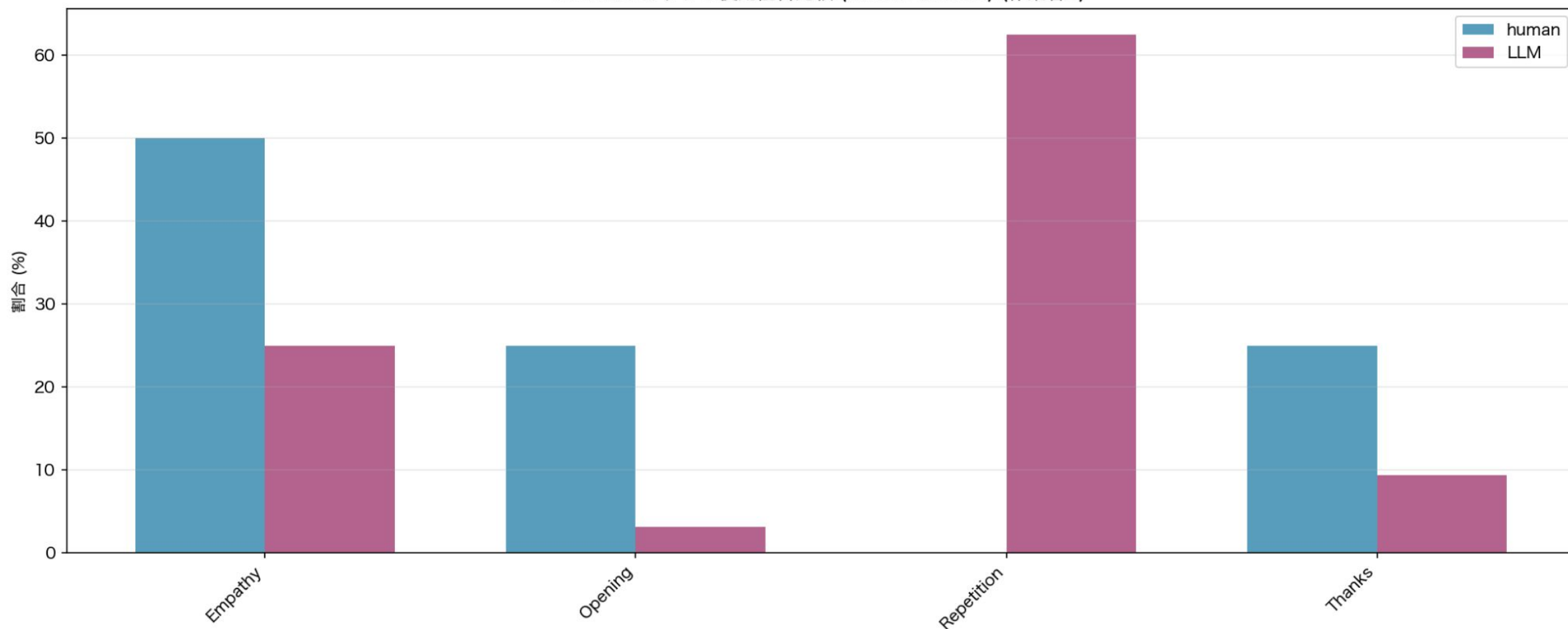
## LLM が General Opinion を多用 人間は Comment を多用

Semantic タグの使用割合比較 (human vs LLM) (作業者B)



## LLM が Repetition を多用 人間は Empathy, Opening, Thanks を多用

Functional タグの使用割合比較 (human vs LLM) (作業者B)



# 考察：MIDASタグセットの課題

- 各タグの包含する概念が広い
  - ISO24617-2のタグをかなり統合
  - MIDASの設計思想
    - タスクが限定された，LLM以前のチャットボットとの対話が対象
    - 人手による大規模アノテーションが目的

MIDAS	ISO
<i>factual question</i>	<i>Question, Set-Question, Choice-Question</i>
<i>opinion question</i>	<i>Question, Set-Question, Choice-Question</i>
<i>yes/no question</i>	<i>Propositional-Question, Check-Question</i>
<i>task command</i>	<i>Instruct, Suggest, Request, Offer, Promise</i>
<i>invalid command</i>	<i>Instruct, Suggest, Request, Offer, Promise</i>
<i>appreciation</i>	<i>AutoPositive, AlloPositive</i>
<i>general opinion</i>	<i>Inform</i>
<i>complaint</i>	<i>AutoNegative, AlloNegative</i>
<i>comment</i>	<i>AutoPositive, AutoNegative, AlloPositive, AlloNegative</i>
<i>statement non_opinion</i>	<i>Inform, Init-Self-Introduction, Return Self-Introduction</i>
<i>other answer</i>	<i>Address-Offer, Address-Request, Address-Suggest, Answer, Correction, Disagreement, Agreement</i>
<i>postive answer</i>	<i>Confirm, Accept-Offer, Accept-Request, Accept-Suggest, Agreement</i>
<i>negative answer</i>	<i>Disconfirm, Decline-Offer, Decline-Suggest, Decline-Request, Disagreement</i>
<i>abandon</i>	
<i>nonsense</i>	
<i>opening</i>	<i>Init-Greeting, Return Greeting</i>
<i>closing</i>	<i>Init-Goodbye, Return Goodbye</i>
<i>hold</i>	<i>Staling, Pausing</i>
<i>thanks</i>	<i>Thanking, Accept Thanking</i>
<i>back-channeling</i>	<i>Propositional-Question</i>
<i>apology</i>	<i>Apology</i>
<i>apology response</i>	<i>Accept Apology</i>
<i>other</i>	<i>Other</i>

ISO24617-2とMIDASのマッピング  
(Yu & Yu 2021: 1120)

# 修正案：新規タグセットの提案 1/2

## パイロットアノテーションを受けてタグセットを修正

- 修正方針
  - MIDASよりもやや詳細なタグが必要
  - 特に、先行発話に対するレスポンスに関わるタグを強化
    - 賞賛
    - 感謝
    - 反復／引用
- 日本語日常会話コーパスへの談話行為アノテーション (居關 et al. 2017, 2018) をベースにISO24617-2 の一部を利用 (cf. [須藤 et al. 2021](#), [近 et al. 2025](#))
  - 音声対話→チャットでの対話に合わせてタグセット・定義を調整

# 修正案：新規タグセットの提案 2/2

## 日本語日常会話コーパスベースの ISO24617-2 互換タグセット

- T\_Agreement / T\_Disagreement, フィードバック系
  - T\_Informに加えて, 賛意や反論に対応
  - 単なる Empathy は漠然としているため細分化
- FB\_Repetition
  - LLM の発話にみられる反復に対応
  - 情報提供系との区別のため
- Qualifier系
  - 感情表現に対応
- 一部は統合・定義を本研究が独自に決定
  - 太字：新たに ISO24617-2 から追加
  - 下線：本研究が独自に定義

分類	タグ
タスク系	T_Inform <b>T_Agreement</b> <b>T_Disagreement</b> T_Question T_CheckQuestion T_Answer T_Request T_A-Request
社会的付き合い管理系	S_Greeting S_Apology S_A-Apology S_Thanking S_A-Thanking
フィードバック系	FB_Auto_Positive FB_Allo_Positive FB_Auto_Negative FB_Allo_Negative <u>FB_Repetition</u>
Qualifier系	<b>Q_Sentiment_Positive</b> <b>Q_Sentiment_Negative</b>

# 新規タグセットによるアノテーションの事例

## 新規タグによるアノテーションが現在進行中

Speaker	message	作業者A	作業者B, C	新規タグ
user	幸福とは <u>自分自身が幸せを感じる時</u> です	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .
. . .	. . .	. . .	. . .	. . .
カナコ (LLM)	カナコ: <u>userさんの「自分自身が幸せを感じる時」という問い、すごく考えさせられますよね。</u> 自分にとっての幸せって、結局は自分自身でしか定義できないものなのかも。 それを見つけるヒントって、どこにあるんでしょうね？	<b>Empathy</b>	<b>Repetition</b>	<b>FB_Repetition</b>

# 課題・今後の予定

- ・ 新規タグによるアノテーションの完遂
  - ・ 10%を人手・LLMでアノテーションし一致率を確認
  - ・ 十分な一致率があれば、残りをLLMでアノテーション
- ・ アノテーション手続きの検討
  - ・ メッセージ単位でのマルチラベルタグ付与が一致度を下げている可能性
- ・ 追加実験によるデータ収集
  - ・ 実験上の指示が LLM の発話に影響
  - ・ LLM のみでの議論をするとどうなるか？
  - ・ 結論が出やすい議論をさせた場合はどうなるか？

\*本研究は、「戦略的イノベーション創造プログラム（SIP）」「統合型ヘルスケアシステムの構築」JPJ012425、および、「次世代研究者挑戦的研究プログラム」JPMJSP2140の支援を受けた。