

Title	発話解釈に用いられる推論形式の実証
Author(s)	松崎, 由幸
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="https://hdl.handle.net/10119/20511">https://hdl.handle.net/10119/20511</a>
Rights	
Description	Supervisor:橋本 敬, 先端科学技術研究科, 修士(知識科学)

修士論文

発話解釈に用いられる推論形式の実証

松崎 由幸

主指導教員 橋本 敬

北陸先端科学技術大学院大学  
先端科学技術専攻  
(知識科学)

令和 08 年 03 月

## Abstract

Utterance interpretation is not completed solely by decoding literal meaning; it involves a process in which the hearer infers what the speaker intends through pragmatic reasoning. While Relevance Theory (Sperber & Wilson, 1986/1995) assumes inferential processes in utterance interpretation, the specific forms of inference are not explicitly specified (Yoshimura, 2016). Meanwhile, Yoshimura (2016) and Hashimoto (2020) have proposed that utterance interpretation is driven by abductive reasoning, but their work remains theoretical and has not been empirically validated. Therefore, the purpose of this study was to empirically investigate which forms of inference hearers employ when deriving interpretations. To achieve this purpose, the following two research questions were formulated:

RQ1: What are the inferential strategies used in utterance interpretation?

RQ2: What individual characteristics are associated with inferential strategies in utterance interpretation?

An experiment was conducted with 38 young adults using six conversational stimuli classified into three levels of comprehension ease (high, medium, and low). Participants provided intuitive interpretations for each stimulus, generated multiple interpretation candidates, and finally selected the optimal interpretation. The interpretation generation process was collected using the think-aloud protocol, and verbal protocols were coded into three inference types: abduction (hypothetical inference of speaker intention from context), rule-based inference (inductive/deductive reasoning based on general knowledge and conventions), and cognitive empathy (inferring intention by adopting the speaker's perspective). Frequency and proportion of use for each inference type were calculated and analyzed in relation to comprehension ease and individual characteristics. Comprehension ease levels were determined based on preliminary survey results.

Two hypotheses were tested. For Hypothesis 1—"The lower the comprehension ease, the greater the use of abduction"—significant differences in abduction frequency were found across stimuli ( $\chi^2(5) = 12.40$ ,  $p = .030$ ). However, no significant differences were observed when comparing representative values by comprehension ease level, and the mean values did not show a monotonic increase according to the expected order (high: AB, medium: CD, low: EF; A: 2.74, B: 2.82, C: 2.16, D: 2.03, E: 1.95, F: 2.61). Therefore, comprehension ease cannot be considered a primary factor that simply increases or decreases abduction use, and Hypothesis 1 was not supported.

For Hypothesis 2—"Individual characteristics related to communication are associated with patterns of inference type use"—communication skills (ENDCOREs: decoding ability and other-acceptance) showed no correlation with abduction frequency. However, they showed significant positive correlations with rule-based inference frequency (decoding ability:  $\rho = .413$ , adjusted  $p = .050$ ; other-acceptance:  $\rho = .446$ , adjusted  $p = .050$ ). Regarding proportions, abduction proportion showed significant negative correlations (decoding ability:  $\rho = -.456$ , adjusted  $p = .029$ ; other-acceptance:  $\rho = -.439$ , adjusted  $p = .029$ ), while rule-based inference proportion showed marginally significant positive correlations (decoding ability:  $\rho = .408$ , adjusted  $p = .055$ ; other-acceptance:  $\rho = .411$ , adjusted  $p = .055$ ). Although measures related to Theory of Mind were also administered, no significant correlations were found. These results suggest that the influence of individual characteristics may manifest not as changes in abduction use, but rather as additional recruitment of inductive/deductive reasoning. Thus, Hypothesis 2 was partially supported.

Additionally, it was shown that the tendency to use different inference types did not change substantially even when task demands varied (intuitive/multiple responses/selection), suggesting that inference type selection may be more strongly determined by individual differences than by task requirements.

The conclusions of this study are threefold. First, abduction appears to be a foundational inference type widely used in utterance interpretation, relatively unaffected by comprehension ease or task demands. Second, individual characteristics related to communication skills are associated with patterns of inference type use, and this influence manifests not as changes in abduction but as additional recruitment of rule-based inference. Third, inference type selection is more strongly influenced by individual differences than by stimulus or task variations, suggesting that inferential strategies in utterance interpretation should be understood as a relatively stable cognitive style rather than a situation-dependent process. These findings suggest that the inferential process in utterance interpretation may have a "hierarchical structure with abduction as the foundation, upon which inductive/deductive reasoning is layered according to individual differences and situational factors."

Three directions for future research are identified. First, the replicability of the finding that rule-based inference is layered upon abduction as a foundational strategy needs to be confirmed with larger samples. Additionally, interaction effects between comprehension ease and individual characteristics (e.g., whether individuals with higher skills show increased rule-based inference for difficult stimuli) should be examined with sufficient statistical power. Second, since this experiment was conducted with Japanese native speakers who were university students, future studies should extend to samples varying in age, cultural background, and language to verify whether similar results are obtained. Third, the stimuli created for this study need to be organized along multiple dimensions (e.g., semantic clarity, knowledge dependency). This would provide further insights into which factors influence the differential use of inference types.

# 目次

第1章 はじめに .....	1
1.1 背景 .....	1
1.2 目的 .....	2
1.3 手法 .....	2
1.4 意義 .....	3
1.5 本論文の構成 .....	3
第2章 関連研究 .....	5
2.1 発話解釈における推論プロセス .....	5
2.1.1 発話解釈とは何か .....	5
2.1.2 関連性理論における発話解釈 .....	5
2.1.3 推論の役割 .....	6
2.2 推論形式の理論的検討 .....	7
2.2.1 推論形式の種類 .....	7
2.2.2 コミュニケーションにおけるアブダクションの構造 .....	8
2.2.3 発話解釈への適用可能性 .....	8
2.3 アブダクションの実証研究 .....	9
2.3.1 非言語領域における実証 .....	9
2.3.2 言語領域における実証 .....	10
2.3.3 先行研究の限界 .....	10
2.4 解釈選択における認知的要因 .....	11
2.4.1 心の理論と認知的共感 .....	11
2.4.2 心の理論個人差 .....	13
2.5 本研究の焦点 .....	13
2.6 発話解釈プロセスの実証的検討 .....	14
2.6.1 推論プロセス測定の難しさ .....	14
2.6.2 Find Out タスク .....	14
2.6.3 Think Aloud 法（思考発話法） .....	15
2.6.4 反応時間 .....	15
2.7 Think Aloud 法による思考の可視化 .....	16
2.7.1 ひらめき学習プロセスにおける Think Aloud 法 .....	16

2.7.2	発話解釈における Think Aloud 法 .....	16
2.7.3	本研究での Think Aloud 法テキスト分析方針 .....	17
2.8	関連研究のまとめと本研究で扱う課題と仮説 .....	17
第 3 章	実験刺激の作成と妥当性検討 .....	19
3.1	目的 .....	19
3.2	実験刺激作成の方針 .....	19
3.2.1	実験刺激の構造 .....	19
3.2.2	統制した要素 .....	20
3.2.3	作成手順 .....	20
3.3	実験刺激の理解容易度分布の確認 .....	21
3.3.1	参加者・手続き .....	21
3.3.2	手順 .....	21
3.3.3	分析方法 .....	22
3.4	結果 .....	23
3.4.1	記述統計量 .....	23
3.4.2	実験刺激の選出 .....	25
3.5	考察 .....	29
3.5.1	理解容易度調査の妥当性 .....	29
3.5.2	本実験への接続 .....	30
第 4 章	推論形式判別基準作成のための予備実験 .....	31
4.1	目的 .....	31
4.2	方法 .....	31
4.2.1	参加者・手続き .....	31
4.2.2	実験刺激 .....	31
4.2.3	手法 .....	32
4.2.4	判別基準の作成方法 .....	33
4.3	結果 .....	34
4.3.1	回答の言語化 .....	34
4.3.2	推論形式の判別結果 .....	34
4.4	考察 .....	35
4.4.1	回答の言語化 .....	35
4.4.2	判別基準の妥当性 .....	35

4.4.3 発話解釈に関する実験への接続 .....	36
第5章 発話解釈に関する実験 .....	38
5.1 参加者についての詳細 .....	38
5.1.1 募集方法 .....	38
5.1.2 参加者と実施期間 .....	38
5.2 実験デザイン .....	39
5.2.1 目的 .....	39
5.2.2 仮説 .....	39
5.2.3 実験計画 .....	39
5.3 実験刺激 .....	<a href="#">3940</a>
5.3.1 実験刺激の内容 .....	<a href="#">3940</a>
5.3.2 刺激の提示順序 .....	40
5.3.3 刺激の提示方法 .....	41
5.4 手順 .....	41
5.4.1 実験の流れ .....	41
5.4.2 インフォームドコンセント .....	42
5.4.3 実験環境 .....	42
5.5 直感回答課題 .....	43
5.5.1 目的 .....	43
5.5.2 手順 .....	<a href="#">4344</a>
5.5.3 収集するデータ .....	44
5.5.4 発話解釈にかかる時間の算出 .....	44
5.5.5 発話解釈にかかる時間の算出対象 .....	45
5.5.6 発話解釈の推論形式の導出 .....	46
5.6 多数回答課題 .....	46
5.6.1 目的 .....	46
5.6.2 手順 .....	47
5.6.3 収集するデータ .....	48
5.6.4 発話解釈の推論形式の導出 .....	48
5.6.5 発話解釈の推論形式の分析対象 .....	48
5.7 解釈選択課題 .....	50
5.7.1 目的 .....	50

5.7.2	手順	50
5.7.3	収集するデータ	51
5.7.4	発話解釈の推論形式の導出	51
5.7.5	発話解釈の推論形式の導出対象	51
5.8	個人特性	51
5.8.1	目的	51
5.8.2	測定した心理測定尺度	51
5.9	半構造化インタビュー	53
5.9.1	目的	53
5.9.2	質問内容	53
5.10	統計手法	53
5.10.1	実験刺激の操作検証	53
5.10.2	仮説の検証	54
5.10.3	心理測定尺度	54
5.10.4	個人特性と推論形式の関係	54
第6章	結果	55
6.1	実験刺激の操作チェック	55
6.1.1	各実験刺激の反応時間分布	55
6.1.2	各実験刺激の反応時間差の検証	56
6.1.3	各刺激のタイムアウト率分布	57
6.1.4	タイムアウト率から見た各刺激の差	58
6.1.5	操作チェックのまとめ	59
6.2	Think Aloud 法で取得した回答の推論形式判別	60
6.2.1	推論形式の判別例	60
6.2.2	推論形式使用回数・比率の傾向	61
6.3	理解容易度が推論形式使用回数に与える影響	63
6.3.1	記述統計量	63
6.3.2	各刺激の推論形式使用回数の差	65
6.3.3	理解容易度の推論形式使用回数の差	65
6.4	理解容易度が推論形式使用比率に与える影響	66
6.4.1	記述統計量	66
6.4.1	各刺激の推論形式使用比率の差の検証	69

6.4.2 理解容易度の推論形式使用比率の差の検証 .....	70
6.5 個人特性が推定形式に与える影響 .....	71
6.5.1 記述統計量 .....	71
6.5.2 個人特性が与える推論形式使用回数への影響 .....	74
6.5.3 個人特性が与える推定形式使用比率への影響 .....	77
6.5.4 使用回数と使用比率の対比 .....	79
6.6 課題間での推論形式の変化 .....	79
6.6.1 各課題の推論形式使用比率の探索 .....	79
6.6.2 各刺激の推論形式使用比率の探索 .....	81
6.6.3 参加者ごとの見た推論形式使用割合 .....	85
6.6.4 課題における推論形式の推移 .....	86
第7章 考察 .....	87
7.1 結果の概観 .....	87
7.2 理解容易度と推論形式の関係 .....	88
7.2.1 仮説の検証結果 .....	88
7.2.2 結果から考えられる発見 .....	89
7.2.3 仮説1 非支持の要因 .....	89
7.3 個人特性と推論形式の関係 .....	90
7.3.1 仮説2 の検証結果 .....	90
7.3.2 演繹追加方略 .....	91
7.3.3 使用回数と使用比率の対比 .....	92
7.4 統合的理解 .....	92
7.4.1 仮説1 と仮説2 の整理 .....	92
7.4.2 仮説1 と仮説2 の統合 .....	<a href="#">9394</a>
7.5 刺激特性の複雑性と理解容易度概念の限界 .....	94
7.6 本研究の限界 .....	94
第8章 おわりに .....	96
8.1 研究の概要 .....	96
8.2 結論 .....	97
8.3 今後の課題 .....	<a href="#">9798</a>

# 目次

図 3.4.1 各実験刺激の理解容易度標定分布 .....	28
図 5.4.1 実験風景 左側が実験実施者のスペース, 右側が参加者のスペース .....	42
図 5.5.1 反応時間(RQ)の算出方法 .....	45
図 5.5.2 発話解釈開始の判別基準.....	45
図 5.6.1 回答分析対象.....	48
図 5.6.2 回答の分析 (推定の場合) .....	49
図 5.6.3 回答の分析 (規則の場合) .....	49
図 5.6.4 回答の分析 (認知的共感の場合) .....	49
図 6.1.1 各刺激の反応時間分布 .....	56
図 6.2.1 各刺激の推論形式使用回数分布 .....	62
図 6.2.2 各刺激の推論形式使用比率分布 .....	62
図 6.3.1 多数回答課題 (Q2) の「推定」使用回数分布 .....	64
図 6.3.2 多数回答課題の「規則」使用回数分布 .....	64
図 6.3.3 多数回答課題の「認知的共感」使用回数分布.....	65
図 6.4.1 多数回答課題の「推定」使用比率分布 .....	68
図 6.4.2 多数回答課題の「規則」使用比率分布 .....	68
図 6.4.3 多数回答課題の「認知的共感」使用比率分布.....	69
図 6.5.1 解読力と各推論形式および総回答数の相関 .....	76
図 6.5.2 他者受容と各推論形式および総回答数の相関.....	76
図 6.5.3 解読力 (end_rT) と各推論形式使用比率の相関.....	78
図 6.5.4 他者受容 (end_aT) と各推論形式使用比率の相関.....	78
図 6.6.1 各課題の推論形式使用比率の推移 .....	80
図 6.6.2 課題の推論形式使用比率の差.....	80
図 6.6.3 各刺激の推論形式使用比率の推移 .....	83
図 6.6.4 各課題の推論形式使用比率の差 .....	84
図 7.3.1 推論形式使用回数・比率と解読力・他者受容の相関.....	91

# 表目次

表 3.4.1	理解容易評定値の記述統計量.....	24
表 3.4.2	理解容易評定値の判別結果.....	26
表 3.4.3	選定された実験刺激の理解度標定.....	27
表 3.4.4	選定された実験刺激の詳細.....	27
表 4.2.1	推論形式の判別基準.....	33
表 4.3.1	追加情報による解釈候補生成数の変化.....	34
表 4.3.2	各推論形式の出現頻度.....	35
表 4.4.1	本分析に使用する推論形式判別基準.....	37
表 5.3.1	実験刺激の詳細.....	40
表 5.3.2	実験刺激の提示と課題の順.....	41
表 5.4.1	実験機器の詳細・配置および役割.....	43
表 6.1.1	各刺激の反応時間の記述統計量.....	55
表 6.1.2	各刺激の反応時間差の検定.....	56
表 6.1.3	理解容易度の差の検証.....	57
表 6.1.4	各刺激のタイムアウト率の記述統計量.....	57
表 6.1.5	刺激別のタイムアウト率分布.....	58
表 6.1.6	各刺激に関するタイムアウト率の比較.....	58
表 6.1.7	理解容易度の差の検証.....	59
表 6.2.1	各刺激の推論形式使用回数・比率全体の傾向.....	62
表 6.3.1	各刺激の推論形式使用回数の記述統計量.....	63
表 6.3.2	各推論形式使用回数の検定結果集計.....	65
表 6.3.3	各刺激の推論使用回数の有意差.....	66
表 6.3.4	各刺激の規則使用回数の有意差.....	66
表 6.4.1	各刺激の推論形式使用比率の記述統計量.....	67
表 6.4.2	各推論形式使用比率のフリードマン検定の集計.....	69
表 6.4.3	各刺激の「推定」使用の差.....	70
表 6.4.4	各刺激の「規則」使用確率の差.....	70
表 6.5.1	心理測定尺度の記述統計量 (N=38).....	72
表 6.5.2	心理測定尺度の信頼性分析 (N=38).....	73
表 6.5.3	個人特性と各推定形式および総回答数の相関.....	75

表 6.5.4 個人特性と「推定」使用比率の相関.....	77
表 6.5.5 個人特性と「規則」使用比率の相関.....	77
表 6.6.1 各課題の推論形式使用比率 .....	81
表 6.6.2 各刺激の各推論形式使用比率.....	82
表 6.6.3 各参加者の Q1・Q3 推論形式一致率 .....	85

# 第1章 はじめに

本研究は、発話解釈における推論形式を実証的に明らかにすることを目的とする。本章では、1.1 節では発話解釈における推論プロセスの理論的背景と未解決の問題を整理する。1.2 節では本研究の目的と研究課題を提示する。1.3 節では本研究の学術的・社会的意義を述べる。1.4 節では本研究の仮説を検証する方法について述べる。最後に 1.5 節では、論文の構成を述べる。

## 1.1 背景

日常会話において、聞き手は話し手の字義的な発話内容だけでなく、その背後にある話し手の意図を手がかりとして推論される意味、すなわち推意 (implicature) を理解することが多い (Grice, 1975)。例として、「窓が開いている」という発話は、文脈によって「窓を閉めてほしい」という依頼や「寒い」という不満として解釈される。このような発話の理解メカニズムは、語用論研究の中心的課題である。

発話理解のメカニズムについて、関連性理論 (Sperber & Wilson, 1986/1995) では推論を進める形式的な規則として演繹的推論を想定する (Sperber & Wilson, 1986/1995:85)。一方で、吉村 (2016) が指摘するように、その推論形式は "non-demonstrative" (非論証的) とされており、具体的にどのような推論形式かは明示されていない。吉村 (2016) は、この二重記述の整合性が明確ではない点を指摘し、発話解釈の導出には演繹だけでなく、帰納やアブダクションを含む拡張的推論形式が関与する可能性を示唆した。

ここで演繹とは、一般的な法則や前提から、個別の事象について必然的に導かれる結論を導出する推論、帰納は、複数の具体的な観察規則から共通点や傾向を抽出し、一般的な法則や規則性を導き出す推論である。そして、アブダクションは、観察された結果からそれを最もよく説明できる仮説を導き出す推論形式であり、限定的な情報からもっともらしい説明を想定的に構築する点が特徴である (Peirce, 1931; 米盛, 2007)。

発話解釈は、観察された発話から話し手の意図という直接観察できない原因を推論するプロセスである。この構造は、観察事象から最良の説明を推論するアブダクションの構造と一致する (米盛, 2007)。したがって、発話解釈においてアブダクションが重要な役割を果たす可能性がある (Hashimoto, 2020)。

発話解釈は話し手の意図を推測するプロセスであり、心の理論(Theory of Mind)が関与する(Premack & Woodruff, 1978). 近年の研究は、推論のプロセスとシミュレーションのプロセスの両方が発話解釈に関与することを示唆している(Ali, 2020). このことは、発話解釈においても複数の推論形式が状況に応じて使い分けられている可能性を示唆する。さらに、心の理論の活用には、個人差があることが示されている。たとえば、Mayn and Demberg(2025)は、他者理解能力が高い参加者ほど語用論的推論において発話解釈が導出しやすいことを示した。

## 1.2 目的

先行研究の理論的検討から、以下の点が示唆される。関連性理論(Sperber & Wilson, 1986/1995)では発話解釈における推論を想定はしているが、具体的な推論形式は明示されていない(吉村, 2016). 一方で、吉村(2016)や Hashimoto(2020)は発話解釈がアブダクションで推論されていると提唱しているが、理論的検討であり実証までは行われていない。そして、発話解釈において推論形式の活用パターンにどのような個人差があるかは知られていない。以上の背景を踏まえ、本研究の目的は、発話解釈においてどのような推論形式をどのように用いているか実証的に明らかにすることである。

具体的には、以下の2つの研究課題に取り組む。

- RQ1 発話の解釈における推論形式の方略はどのようなものか
- RQ2 発話の解釈における推論形式の方略には、どのような個人特性が関係しているか

本研究では、それぞれの課題に対して以下の仮説を設定する。この仮説の導出は2章で行う。

- 仮説 1 発話が困難なほどアブダクションが用いられる
- 仮説 2 コミュニケーションと思考特性に関する個人特性は、推論形式の使用頻度と関連する

## 1.3 手法

本研究では、2人の短い会話テキストを提示し、一方の発話者(返答者)の意図を推測してもらう課題を実施した。参加者に推論を要求するため、会話テキストは返答が理解しやすさ(以降、理解容易度)を操作した、理解容易度が高いものから低いもの6種類を用意した。課題は、短時間で直感的に返答者の意図を

答える「直感回答」、Think Aloud 法を用いてできるだけ多数の解釈を生成する「多数回答」、その中から適切だと思う1つを選択させる「選択」という3つを実施した。Think Aloud のデータから用いられている推論形式を同定し、それが理解容易度に応じてどのように変化するかを分析した。また、個人特性を調査するために上記の課題実施後に心理測定尺度を測定し、用いられた推論形式との関係を分析した。

会話テキスト作成の詳細を第3章、予備実験を通じた推論形式判別基準の作成の詳細を第4章、3つの課題および心理測定尺度の詳細を第5章にそれぞれ示す。

## 1.4 意義

本研究の学術的意義は以下の二点にある。一つ目は、発話解釈の推論形式について、関連性理論が抽象的に記述してきた「推論的理解」の内実を具体化し、語用論の精緻化に貢献できる。二つ目は、語用論研究と認知心理学を橋渡しする基礎的知見を提供する点である。Mayn and Demberg(2025)が指摘するように、語用論的推論における個人差の要因を調べた研究は少なく、推論形式と個人特性の関係は十分に検討されていない。本研究でこれらの関係を検討することで、発話解釈における個人差がどのような推論方略の違いに起因するのか明らかになる。

社会的意義は次の二点である。一つ目は、教育やコミュニケーション支援への応用である。本研究から得られる知見は、発話解釈が苦手・困難な人々に対し、どの段階でどのような支援が有効であるかを検討する基礎的知見となる。二つ目は、対話システムの自然言語理解の改善である。人間が発話解釈にどのような推論形式をどのように用いるかを明らかにすることで、チャットボットでの会話や対話型生成 AI にユーザーが与えるプロンプトから、ユーザーが持つ意図を推測する技術開発の基礎になると考える。

## 1.5 本論文の構成

本論文は全8章で構成される。第2章では、発話解釈における推論プロセスと推論形式に関する先行研究を整理し、本研究の位置づけを明確にした上で、1.2節で示したRQと仮説を導出する。第3章では、意図の推測しやすさを統制した実験刺激を作成するため、オンラインアンケートを実施して意図の推測しやすさのデータを収集し、刺激の選定した詳細を述べる。第4章では、Think Aloud

法によって得られた回答から推論形式を判別する基準を確立する。第5章では、本実験の詳細（参加者・手続き・課題の詳細・測定指標・分析方針）を述べる。第6章では、実験の結果について述べ、仮説検証のための分析を行う。第7章では、第6章の結果を総合し、理解容易度・個人特性・課題（直感／多数回答／選択）の観点から推論形式の使用パターンを考察する。その上で、仮説1と仮説2の統合的解釈を提示し、理解容易度概念の限界と本研究の限界を述べる。第8章では、本研究で得られた知見を要約し、結論や本研究の限界と今後の課題を述べる。

## 第2章 関連研究

本章では、発話解釈における推論形式をどのように用いているかを明らかにするための理論的・実証的基盤を整理し、本研究で扱う具体的研究課題 (Research Question, RQ) とそれに対する仮説を導出する。2.1 節では発話解釈における推論プロセスの基本的性質を検討し、推論形式の同定という課題の重要性を示す。2.2 節では推論形式を整理し、アブダクションが発話解釈の推論形式として有力な候補であることを理論的に検討する。2.3 節ではアブダクション推論の実証研究を概観し、発話解釈への適用可能性を検証する。2.4 節では心の理論や認知的共感といった個人差要因が解釈選択にどのように関与するかを論じる。2.5 節では発話解釈プロセスを実証的に研究するための方法論を検討する。2.7 節では本章のまとめと本研究の二つの研究課題、それぞれに対応する仮説を提示する。

### 2.1 発話解釈における推論プロセス

#### 2.1.1 発話解釈とは何か

発話解釈は、字義的意味の復号に加え、文脈依存的な推意を推論的に構成する過程である (Grice, 1975; Sperber & Wilson, 1986/1995)。例として「窓が開いている」という発話は、窓の物理的状态を記述する字義的意味を持つが、文脈次第で「窓を閉めてほしい」という依頼や「寒い」という不満として理解される。このように、同一の発話であっても、状況・話し手と聞き手の関係性・共有知識によって、複数の異なる解釈が成立しうる。

この解釈の多義性こそが、発話解釈を単なる言語的デコーディングではなく、推論を必要とする認知的プロセスとして特徴づけている。聞き手は、発話という観察可能な情報から、話し手の意図という直接観察不可能な心的状態を推測しなければならない。この意味で、発話解釈は不確実性下での仮説形成と選択を本質的に含む認知過程として捉えられる。整理すると、発話解釈とは (1) 不確定性をもち、(2) 複数解釈候補の存在があり、(3) 話し手の意図を説明・理解しようとするものといえる。

#### 2.1.2 関連性理論における発話解釈

Sperber and Wilson (1986/1995) の関連性理論は、この推論的プロセスを体系的に説明する代表的な理論である。同理論では、聞き手は発話が「十分な認知効果を、過度な処理努力なしに得られる」ことを前提とし、文脈情報と発話内容を統

合して解釈を構成する。ここで認知効果とは、新しい情報が既存の知識体系に統合されることで得られる認知的変化を指し、処理労力とは、その統合に要する認知資源の量を指す。聞き手は、文脈情報と発話内容を理解しようとする労力とのバランスが最適化された解釈（最も関連性の高い解釈）を選択すると想定される（Sperber & Wilson, 1986/1995）。

### 2.1.3 推論の役割

以上より、発話解釈において推論は不可欠な役割を果たすと考えて良いだろう。しかし、理論的課題は「発話解釈が推論される」という一般論ではなく、どのような推論形式・操作によって解釈仮説が生成され、どの基準で選択されるのかという点にある。関連性理論は推論を中核概念として据える一方で、仮説生成・選択の具体的メカニズムについては抽象的に記述されやすい。

吉村（2016）によれば、Sperber and Wilson（1986/1995）は発話解釈が演繹的な推論規則の使用を想定する一方で、「the process of inferential comprehension is non-demonstrative（非論証的）」（Sperber & Wilson, 1986/1995:65）とも述べており、「この 2 つの記述がどのように整合するのか、よくわからない」（吉村, 2016:209）と指摘している。ここでの非論証的とは、前提が真だとしても結論が必ず真になるとは限らない推論のことを示すが、その具体的な推論形式は明示されていない。例として、「窓が開いているね」という発言から「窓を閉めて欲しい」という依頼の意図だと推論したとしても、話し手は単に事実を述べただけの可能性や部屋の換気をしてることを評価している可能性もあり、同じ前提から複数の異なる解釈が導かれる。このように、発話解釈においてどのような推論形式がどう用いられて発話理解に至っているかはまだ明らかとはいえず、実証的に明らかにする必要がある。

発話解釈における推論形式の使用方略を明らかにするためには、推論形式自体の機能について、どの推論形式が発話解釈でどう働くと考えられるかを理論的に整理する必要がある。具体的には、不確実性、複数候補の存在、話し手の意図を説明・理解する必要性といった発話解釈の特性にどの推論形式がどう適合するかを検討するのが有効である。次節では、推論形式の類型を整理し、発話解釈に最も適した形式を検討する。

## 2.2 推論形式の理論的検討

### 2.2.1 推論形式の種類

論理学において、推論は大きく帰納、演繹、アブダクションの三つに分類される。本節ではこの三つの推論形式について概説する。

帰納推論は、個別の観察規則から一般的な規則を導く推論形式である。例として「過去に何度も、窓が開いていると言った人は窓を閉めてほしがっていた。ゆえに、窓が開いていると言う人は一般的に窓を閉めてほしがっている」という推論が帰納に該当する。帰納推論は経験に基づく一般化を可能にするが、結論の真偽は確実ではなく、蓋然的である。発話解釈においても、過去の類似した発話経験から解釈を導くという点で帰納的側面があるが、個々の発話は文脈依存性が高く、単純な一般化では説明が不足しやすい。

演繹推論は、一般的な規則から特定の結論を必然的に導く推論形式である。例として「窓が開いていれば部屋は寒くなる。窓が開いている。ゆえに部屋は寒くなる」という推論が演繹に該当する。この推論形式の特徴は、前提が真であれば結論も必然的に真となる論理的確実性にある。しかし、発話解釈においては、前提となる規則や前提集合が常に明示されているわけではない。したがって、演繹だけで発話解釈の全過程を記述することは困難である。

これら二つに対し、アブダクション推論は、観察された事象に対する最良の説明を仮説として推論する形式である。Peirce(1931)によれば、アブダクションは「驚くべき事実 C が観察された。もし A が真ならば C は当然である。ゆえに、A が真であると考えられる理由がある」(Peirce, 1931:188)という形式を持つ。例として、「窓が開いている」という発話(C)が観察されたとき、「もし話し手が窓を閉めてほしいと思っている(A)ならば、『窓が開いている』と言うのは当然である」と推論するのがアブダクションである。この推論は、結論の真偽を保証するものではなく、観察事象を最もよく説明する仮説を生成する。重要なのは、演繹や帰納と異なり、アブダクションは仮説を創造的に生成する点にある。

米盛(2007)は Peirce のアブダクション概念を整理し、それが仮説形成と仮説選択という二段階から成ることを論じた(米盛, 2007:68)。第一段階の仮説形成では、観察事象を説明する複数の仮説候補が生成される。第二段階の仮説選択では、それらの中から「最良」の説明を選択する。

### 2.2.2 コミュニケーションにおけるアブダクションの構造

次にコミュニケーションにおけるアブダクションをどのように捉えることができるのかを整理する。吉村(2016)は、発話解釈における推意導出が演繹だけでなく、アブダクションや帰納といった多様な推論形式によって行われる可能性を理論的に論じている。また、推意を「演繹のような分析的推論による分析的推意(演繹される推意)と、帰納やアブダクションのような拡張的推論による拡張的推意(創作される推意)」(吉村, 2016:215)に分類が可能であることを具体的な発話事例を通じて示した。

Hashimoto(2020)は、アブダクションにおける仮説生成のメカニズムとして、再帰的結合が関係していることを指摘した。再帰的結合とは、複数の概念要素を段階的に組み合わせて、より複雑な階層的構造を構築する認知操作である。この操作は言語における統語構造の形成だけでなく、概念形成や問題解決においても機能する。

また、仮説選択の段階において身体化シミュレーションが重要な役割を果たす統合仮説を提唱している。身体化シミュレーションとは、他者の行動や意図を理解する際に、自分自身がその行動を行う場合の身体的・感覚的経験をシミュレートする認知過程である。Hashimoto(2020)が提唱する理論的仮説では、アブダクションにおいて、再帰的結合によって多様な仮説が生成された後、それらの仮説は身体化シミュレーションによって評価される。この統合仮説は、発話解釈における仮説生成と選択の両段階に、それぞれ異なる認知メカニズムが関与することを示唆している。

### 2.2.3 発話解釈への適用可能性

本章での理論的検討を踏まえると、アブダクションは発話の解釈を導く認知的過程に含まれる推論形式として最も有力な候補であると考えられる。その理由は二つある。第一に、Peirce(1931)の定式化したアブダクションの推論構造は、観察可能な発話から観察不可能な意図を推論する過程に対応する。第二に、米盛(2007)が整理した二段階構造(仮説形成と仮説選択)は、複数の解釈候補の生成と選択という発話理解の過程に沿っている可能性がある。

また、コミュニケーションにおけるアブダクションの理論的枠組みを、関連性理論と統合することで、発話解釈の認知的メカニズムを明らかにする可能性を示唆する。吉村(2016)も提唱していた Sperber & Wilson (1986/1995)の推論的理解過程と関連性の原則をアブダクションの二段階として再解釈し、新たに

Hashimoto(2020)の統合仮説(再帰的結合による仮説生成と身体化シミュレーションによる仮説選択)を適用することで、発話解釈がアブダクションで行われていることの理論的説明が可能になると考える。

ただし、この理論が実際の認知過程を反映していることを示せていない。吉村(2016)は、具体的な事例に基づく示唆に留まっている。本研究では、この理論が実際の認知過程ではどのようになっているかを明らかにする、次節では、アブダクションに関する実証研究を概観し、この理論が実際の発話解釈過程をどの程度説明しうるかを検討する。

## 2.3 アブダクションの実証研究

### 2.3.1 非言語領域における実証

アブダクションの実証研究は、主に自然現象の理解や問題解決場面において蓄積されてきた。その代表的な例が、ブラックボックス課題を用いた研究である。この課題では、参加者は箱の中に隠された原子の位置を、光線の入射位置と出射位置の観察から推論する。観察された光線のパターンから、原子の位置に関する仮説を生成し、すべての観察を説明できる最も単純な説明(原子の数が最小のもの)を選択する(Klichowicz et al., 2020)。

Klichowicz et al.(2020)は、視線追跡(eye-tracking)を用いてブラックボックス課題におけるアブダクションのプロセスを詳細に検討した。この研究では、参加者は複数の光線パターン(観察)を逐次的に提示され、各観察に対して原子配置の仮説を生成することが求められた。視線追跡の結果、参加者が新たな観察を解釈する際に、過去に生成した原子配置の位置に視線を向けることが明らかになった。これは、参加者が記憶から過去の仮説を検索し、それを新たな推論の基盤としていることを示唆する。

彼らの研究で特に重要な知見は、原子配置が光線パターンよりも多く注視されるという点である。この結果は、アブダクションが逐次的に行われ、各段階で構築された仮説が次の段階の推論基盤となることを示す(Klichowicz et al., 2020:3120)。さらに、視覚的に提示された情報は「外部記憶」として機能し、記憶からの検索の必要性を軽減することが示された(Klichowicz et al., 2020:3129)。ただし、作業記憶への負荷が高い場合、参加者はすべての観察を統合的に説明する一貫した仮説体系を構築するのではなく、個々の観察を独立に説明する傾向があることも明らかになった(Klichowicz et al., 2020:3129)。

この知見を発話解釈に置き換えると、先行する解釈が後続する発話解釈に影響を与える可能性があることを示唆する。すなわち、対話の時間的展開において、先行する解釈候補が後続の推論の基盤となり、また認知的負荷によって統合的な解釈方略と局所的な解釈方略が使い分けられる可能性がある。

### 2.3.2 言語領域における実証

言語理解におけるアブダクションの研究は、非言語領域に比べて限られているが、近年いくつかの重要な知見が報告されている。秋元・邑本(2007)は、皮肉理解における推論プロセスを反応時間と認知的負荷の操作を用いて検討した。皮肉表現（例：大雨の中で「いい天気だね」）と言う場合、聞き手は字義的意味と文脈の矛盾を解消する推論が必要となる。

秋元・邑本(2007)は、共通で持つ情報の有無によって皮肉らしさの知覚が変化することを示した。ここでの共通で持つ情報とは、話し手と聞き手の間で共有されている知識や前提を指す。秋元・邑本(2007)の研究で特に重要な知見は、認知的負荷がかかった条件では、聞き手は話し手と共通で持つ情報を考慮できず、自己中心的視点から語用論的不誠実を検知した時点で発話のアイロニーらしさを知覚することが明らかになった点である。語用論的不誠実とは、「発話行為の適切性条件に違反する発話であり、事実と反するものに加えて、発話の表面的な形式が発話の意図に誠実でないもの」(秋元・邑本, 2007:292)である。例として、実際には窓が閉まっているのに「窓が開いている」という場合や、皮肉として「今日はいい天気ですね（実際には雨）」という場合を示す。

この知見は、皮肉理解においてもアブダクションのような仮説生成プロセスが機能していること可能性を示す。これは、自分が考えられる労力が十分に余っている場合、聞き手は話し手との共通で持つ情報を考慮して「話し手はこの発話を皮肉として意図している」という仮説を立てるが、労力が無く限界の場合にはこの生成プロセスを省略し、自己中心的な解釈に基づいて皮肉を知覚している可能性が考えられる。

### 2.3.3 先行研究の限界

以上の実証研究の検討から、発話解釈における推論形式の解明には三つの課題が残されていると言えるだろう。第一に、発話解釈におけるアブダクションの実証的検討が十分ではない。非言語領域、特にブラックボックス課題のような物理的推論場面ではアブダクションが実証されているが、皮肉理解のような特殊

な言語現象では一部検討されており、日常的な発話解釈における推論形式の同定は依然として課題である。

第二に、複数解釈の生成プロセスについて実証的な解明が不十分である。吉村(2016)や Hashimoto(2020)は理論的に示したが、それらがどのような順序で、どのような基準で生成されるかについては明らかでない。

第三に、解釈選択のメカニズムが不明瞭である。関連性理論は「最も関連性の高い解釈が選ばれる」(Sperber & Wilson, 1986/1995)と述べるが、その選択プロセスに他の認知的要因、例えば話し手の心的状態を理解する能力(心の理論)や他者視点を構成する能力(認知的共感)が関与するかは検討されていない。これらの認知的要因は、特に解釈選択の段階において重要な役割を果たす可能性があるが、その実証的検討は不足している。

これら三つの限界は相互に関連していると考えられる。推論形式を同定するには、複数解釈の生成プロセスを観察する必要があり、選択メカニズムを理解するには、生成された複数の候補の中からどのように一つが選ばれるかを検討する必要がある。したがって、これら三つの課題を統合的に扱う実証研究が求められる。特に解釈選択メカニズムを解明するには、話し手の意図を推測する能力、すなわち心の理論や認知的共感といった認知的要因を考慮した実証的検討が必要である。次節では、これらの概念と発話解釈の関係を検討する。

## 2.4 解釈選択における認知的要因

### 2.4.1 心の理論と認知的共感

心の理論(Theory of Mind)とは、他者が自分とは異なる信念、欲求、意図を持つことを理解する能力である(Premack & Woodruff, 1978)。心の理論に関する研究は当初、チンパンジーが他者の心的状態を推測できるかという問いから始まったが、その後、人間の発達や社会的認知の中核的能力として広く研究されるようになった。

心の理論は発話解釈において中心的な役割を果たすと考えられている。発話解釈とは本質的に話し手の「意図」を推測する行為だからである。例として「窓を閉めてほしい」という意図を「窓が開いている」という発話から理解するには、話し手の欲求や目標を推測する能力が不可欠である。

心の理論をめぐるのは、理論説、シミュレーション説、そして両方を統合したハイブリッド説が提案されてきた(Ali, 2020)。理論説では、心の働きについて信

念や欲求といった一般的なルールを用いて推論をすることで相手を理解すると考えられている。シミュレーション説では、自分を相手の立場や目線に置き、自分の経験を基に他者を理解すると考えられている。ハイブリッド説では、知識に基づく推論とシミュレーションを併用して相手を理解すると考えられている。神経科学では、脳の一部が他者の心的状態の内容の推論に関与していることが示された(Saxe & Kanwisher, 2003)。ただし、Ali(2020)は発話解釈に関して、視覚や聴覚から得る相手の情報だけでは、相手の言いたいことを一つには決める(選択)ことはできず、そのため、状況や目的といった情報も組み合わせて、意味を絞り込む必要があると示唆している。本研究では、この協働関係をアブダクション(仮説生成・選択)という観点から捉え直し、両プロセスが段階に応じて異なる役割を担う可能性を検討する。

Zambrano & Suárez (2020)は、スカラー含意の処理において心の理論がどのように影響するかを実証的に検討した。スカラー含意とは、「some (いくつか)」という量化詞が「not all (すべてではない)」と解釈される推意現象である。例として「クッキーが何枚か残っている」という発話は、字義的には「少なくとも数枚は残っている」という意味だが、推意として「全部は残っていない(=何枚かは食べられている)」という解釈が導かれる。Zambrano and Suárez (2020)は、登場人物の信念や意図を含む文脈を提示した後に提示文が正しいか間違っているかの判定をさせる課題を実施した。例として「真：いくつかの鳥は鷺である、偽：トラの中には魚もいる」(Zambrano & Suárez, 2020)を参加者に真偽を判別させた。その結果、判別が正しい試行に限り登場人物の信念や意図の有無によって反応時間に統計的な差が認められた。この結果は、登場人物の信念や意図の記載の有無で、スカラー含意を含む文の真偽判定に要する時間が変わる可能性を示した。したがって、発話の意図が不明瞭なテキストであるほど、参加者は話し手の信念や意図の解釈を生成し、推論形式の選択や反応時間に変化が起きると考えられる。

認知的共感(cognitive empathy)は、他者の視点や心的状態を理解する能力であり、心の理論と概念的に近接している。情動的共感とは、他者の感情を自分も感じる能力であるのに対し、認知的共感とは他者の視点を理解する能力である。例として、悲しんでいる友人を見て自分も「悲しい」と感じるのが情動的共感であり、「なぜ悲しんでいるのか」を推測するのが認知的共感である(Shamay-Tsoory et al., 2009)。

心の理論と認知的共感の関係については、研究者間で見解が分かれている。Shamay-Tsoory et al., (2009)は両者は概念的に近く、部分的に重なり合う能力として考えている（例：認知的共感を視点取得・心の理論と関連づける）。本研究でも、発話解釈という観点からは、他者の心的状態を推測することと他者の視点を取得することはどちらも解釈生成・選択に関与すると考える。

ただし、本研究では2つを完全に同一視するのではなく、心の理論を「個人差としての他者理解能力」、認知的共感を「話者の視点に立って意図を推測する推論」として区別して扱う。具体的には心の理論に相当する側面は心理測定尺度によって個人特性として扱う。認知的共感に相当する側面は、Think Aloudに現れる推論操作として、推論形式の分析に含める。（例として、情動語など相手の立場に立ち解釈を組み立てる推論を、本研究では便宜上「認知的共感」と呼ぶ）。また、このような他者理解の働きには、日像的な対話の運用とも関係すると考えられるため、本研究では対人コミュニケーションにも着目する。発話解釈では、発話の表層情報だけでなく、相手の立場や意図をどの程度読み取れるか、どの程度相手の視点を受け止められるかといった、対人情報への処理が推論形式に影響があると考えられる。

#### 2.4.2 心の理論個人差

心の理論や認知的共感には個人差があり、Gallistl et al.(2024)は、愛着が他者の意図や視点を推測できる能力と正の相関を示し、他者の心的状態を推測する能力が個人によって異なることを実証した。また、Mayn and Demberg (2025)では、語用論的推論において他者の意図や視点を推測できる能力が高い参加者ほど、推意を導出しやすいことを示した。この知見は、他者の意図や視点を推測する能力には個人差が影響する可能性を示唆する。

こうした能力の差が発話解釈における推論形式とどのような関係があるのかという知見、特に仮説生成と選択という段階を区別した形での十分な検討は既存研究では限定的である。発話解釈がアブダクションを含む過程であると仮定すると、他者の心的状態を推測する能力は仮説を生成する時と、解釈候補から選択する時で異なる形で関与している可能性がある。

### 2.5 本研究の焦点

以上より本研究の焦点は二点に整理される。一つ目は、発話解釈がどのような

推論形式で行われているのか明らかにすることである。二つ目は、心の理論やコミュニケーション・スキルに関連する個人特性が、発話解釈における推論形式の使用パターンとどのように関係しているのか明らかにすることである。これにより、発話解釈における推論メカニズムの理解に近づくことが可能になる。次節では、このような課題に取り組む関連研究を紹介しながら、本研究の目的達成のための研究方法論を検討する。

## 2.6 発話解釈プロセスの実証的検討

### 2.6.1 推論プロセス測定の難しさ

発話解釈における推論プロセスを実証的に検討する上で最大の困難は、それが内的な認知プロセスであり、直接観察が困難である点にある。Zu and Xia(2012)は、従来の研究では解釈結果を質問紙等で測定する手続きが主に用いられてきた。一方で、質問紙法では発話解釈中に参加者が実際に何を考えているか捉えにくいとも指摘している。

### 2.6.2 Find Out タスク

Żelechowska et al.(2020)が開発した Find Out 課題は、アブダクションにおける仮説生成と仮説評価の両プロセスを観察可能にする新しい方法論である。この課題は、三つの段階から構成される。第一段階のウォームアップ段階では、参加者は曖昧な状況を説明する仮説を自由に生成する。第二段階の協力段階では、参加者は生成した仮説を実験実施者に yes-no 質問することによってテストし、新たな情報に基づいて仮説を更新する。第三段階の最終報告段階では、最も妥当な説明とその根拠を提示する。

この課題の利点は、仮説生成だけでなく仮説評価のプロセスも捉えられる点にある。Żelechowska et al. (2020)は、参加者が質問を通じて段階的に情報を収集し、仮説を洗練させていく過程を詳細に記録できることを示した。さらに、各質問には正当化が求められるため、参加者の推論過程の根拠も明らかになる。加えて、この課題は定性的分析と定量的分析の両方を可能にし、アブダクションの個人差を多角的に検討できる点で、従来の方法にはない利点を持つ。

Find Out 課題では、参加者の推論した内容を記述式で回答する形式を採用しているが、本研究でそのまま適用することは困難である。なぜなら、次の二つの問題点があるからである。(1) 推論と同時にそれを言語化するという二重の認知

的負荷がかかる、(2) 記述の手間を避けるために意図的に回答を制限しようとする可能性がある。これらの問題により、記述式で観察する推論プロセスが参加者の自然な推論プロセスを正確に反映しているとは限らない。したがって、この課題で観察されるプロセスが、自然な発話解釈を正確に反映しているかは慎重に検討する必要がある。

### 2.6.3 Think Aloud 法 (思考発話法)

そこで本研究は Think Aloud 法(思考発話法)を採用することにした。Think Aloud 法は、参加者が課題を遂行しながら思考内容を声に出して報告する方法であり、認知プロセスのリアルタイム測定を可能にする。この方法の利点は、推論プロセスが進行中に言語化されるため、事後報告の限界を克服できる点にある。参加者の発話プロトコルから、推論の過程で何が考慮され、どのような情報が統合されているかを詳細に分析できる。

しかし、Think Aloud 法にも言語化の問題という限界がある。すべての認知プロセスが言語化できるわけではなく、特に無意識で処理していることは報告されにくい。発話解釈の一部、特に高度に定型化された推意の処理は自動化されている可能性があり、その場合、Think Aloud 法では捉えられない。本研究では、言語化を考慮しつつ、Think Aloud 法で推論プロセスの探索的研究を実施する。分析方法の方針は 2.6 節で述べる。

### 2.6.4 反応時間

反応時間は、課題処理に要した時間を表す指標であり、発話解釈に伴う処理負荷や迷いの大きさを間接的に捉える手がかりとなる。ただし、反応時間は推論そのものだけでなく、読解処理や回答するのに言語化する処理など複数の要因にも影響されるため、注意は必要である。

秋元・邑本(2007)は、皮肉理解課題において、数秒程度の時間制約下でも判断が可能であることが報告しており、本実験の課題で時間制約を設ける際の参考となる。ただし課題構造が異なるため、時間をそのまま適用することは避けた。

Zambrano and Suárez (2020)は、画面に出てくる文が正しいかどうかを、真偽で素早く答える課題において、登場人物の心的状態を含む短い分と含まない文で比較している。その結果、回答の正誤を区別せず、全試行を含めた反応時間では心の状態に関する文を見せたことによる影響は有意ではなかった。一方で、正答のみの反応時間に限定して分析すると有意に速くなることを示した。つまり、登

場人物の気持ちや考えを意識させる文を先に提示すると、その後の判断が平均的に速くなり、答える速さの個人差も小さくなる傾向が示された。これは、心的状態に関する文脈操作が推意処理に影響しうる可能性を示唆する。

ただし、反応時間のみでは推論形式のいずれが用いられているかを区別できない。反応時間は処理の複雑さや認知負荷を反映するが、どのような推論方略が行われているかは特定できない。例として、反応時間の延長は、単に文の理解が難しいためかもしれない。したがって本研究では、反応時間測定を実験刺激の独立性調査として使用して、推論プロセスをより直接に捉える方法と併用することで発話解釈における推論プロセスを多面的に検討する。本研究での反応時間の算出についての詳細は 5.5.4 節で述べる。

## 2.7 Think Aloud 法による思考の可視化

本研究では推論プロセスをより直接に捉える方法として Think Aloud 法を採用する。本節では、Think Aloud 法についての先行研究から推論形式がどのようにすると観察可能となるか検討する。

### 2.7.1 ひらめき学習プロセスにおける Think Aloud 法

Ash et al.(2012)は、洞察問題（マッチ棒や硬貨を使った配置パズル）と算数問題という 2 種類の課題を実施した。参加者には各問題を解く際に何を考えているか声に出して話すよう求めた。各問題には最大 4 分間の制限時間が設けられた。分析方法は、次の方法を用いた。1つ目は行き詰まり（黙り込む、わからないという発言）を、独立した評価者が録画映像を見ながら確認した。2つ目は、問題を読み始めてから答えを出すまでの時間を計測した。更に一週間後に同じ問題を回答させ、時間変化を確認した。3つ目は、問題の理解度の評価である。問題を解く前に、問題の各要素について解決にどの程度重要かを評価してもらい、評価したものが解決に必要な要素と一致しているかを分析した。この研究では、発話内容の詳細な分析よりも、行き詰まりの有無や解決時間といった客観的な指標を中心に Think Aloud 法を用いていた。

### 2.7.2 発話解釈における Think Aloud 法

Zu and Xia(2012)は、橋渡し参照と呼ばれる言語現象の理解プロセスを明らかにするために Think Aloud 法を使用した。橋渡し参照とは、文章中で名詞されていない対象を推論し理解する現象である(Zu & Xia, 2012)。例として、「クニック

用の食べ物を買いました。ビールは温かかったです」という文では、「ビール」は最初の文では明示されていないが、食べ物の一部として推論する必要がある (Zu & Xia, 2012;119)。課題は 22 個の英文と橋渡し参照の指示対象を特定させる質問を実施し、回答に Think Aloud 法を用いた。

分析は以下の方法が用いられた。1 つ目は思考過程の分類である。参加者の発話を分析し、どのような情報を使って答えを選んだかを分類した。具体的には、因果関係、時間の順序、一般知識などといった観点のどれに関連するか分類した。2 つ目は逐語的な文字起こしである。「えーっと」「あー」などの言いよどみもすべて記録し、沈黙の長さも「…」の数で表現した。特に 5 秒以上の沈黙は重要な指標として記録した。3 つ目は、振り返りインタビューである。Think Aloud の直後に、録音を一緒に聞きながら、不明確な部分について質問し、参加者の考えを確認した。4 つ目は、問題に直面した兆候の特定である。明確な問題の表明、不明な箇所への下線、意味の分析、ためらい、答えへの評価、話し方の変化といった 6 つの兆候に注目した。

### 2.7.3 本研究での Think Aloud 法テキスト分析方針

以上の 2 つの研究を比較すると、Think Aloud 法の分析方法に明確な違いがあることがわかる。Ash et al. (2012) は、行き詰まりや解決時間といった客観的な指標を中心に分析を行った。一方、Zu and Xia (2012) は、発話内容を詳細に分類・分析し、思考プロセスそのものを明らかにしようとした。このように、Think Aloud 法は研究の目的に応じて、異なる分析方法を用いることができる柔軟な手法であると言える。

本研究では、Zu and Xia (2012) の手法を参考に分析を行う。具体的には、参加者に発話の解釈を自由に考えさせ、その思考過程を声に出してもらい録音する。その録音から文字起こしを行い、参加者がどのような推論形式を用いたかを判別する。そのためには、推論形式を判別する基準が必要である。本研究では、予備実験として推論形式の判別基準を作成する (第 4 章)。また、回答直後のインタビューは解釈選択のプロセスを観測する上で最も観察可能性がある。

## 2.8 関連研究のまとめと本研究で扱う課題と仮説

本章では、発話解釈を「観察可能な発話から話し手の意図を推測する解釈生成・選択の過程」として捉える理論的かつ実証的知見を整理した。理論的観点か

らはアブダクションが有力な推論形式となり得ること、また共有知識や認知的負荷、話し手の心的状態を明らかにする手がかりが推論形式に影響しうることを確認した。

以上を踏まえ、本研究では(1)発話解釈で用いられる推論形式を観察し、(2)心の理論やコミュニケーションに関する個人特性が、解釈の生成にどのように関連しているかを検討する。2.1～2.4節にまとめた先行研究から、本研究ではそれぞれに対応した次の仮説を立てる。

- 仮説1 発話が困難なほどアブダクションが用いられる
- 仮説2 コミュニケーションおよび思考特性に関する個人特性は、推論形式の使用頻度と関連する

これらの仮説を検証するための方法として、第3章で理解容易度を統制した実験刺激の作成および検証、第4章で推論形式判別基準を確立する。そして第5章の本実験では、直感回答、Think Aloud法による多数回答・その中から最も適切な解釈を1つ選択させる。これら3つの課題により、解釈生成と選択を分離して測定し、参加者の回答による推論形式の判別、反応時間、心理測定尺度を併用して発話解釈における推論形式を多角的に捉える。

## 第3章 実験刺激の作成と妥当性検討

第2章では、発話解釈における推論形式をどのように用いているかを明らかにするための理論的・実証的基盤を整理した。本研究では吉村(2016)やHashimoto(2020)は発話解釈がアブダクションで推論されている理論的検討が、実際の発話解釈ではどのような推論形式であるかを明らかにする。

第3章では、実験で理解容易度を統制した会話テキスト（以降、実験刺激）作成の詳細を述べる。3.1節では、実験刺激を作成する目的について述べる。3.2節では、実験刺激作成の方針について述べ、作成手順の詳細を示す。3.3節では、作成した会話刺激（本稿では選出前の会話テキストを“会話刺激”として識別する）の理解容易度評価値（本稿では選出前の理解容易度を“理解容易度評価値”として識別する）についてオンラインアンケートで評価を収集し、評価値の分布を確認した上で、得点に基づいて刺激を高・中・低に分類し、本研究で使用する刺激を示す。3.4節では、調査の結果について述べる。3.5節では、結果についての考察を述べ、選定した実験刺激を示す。

### 3.1 目的

本研究は、発話解釈においてどのような推論形式をどのように用いているか実証的に明らかにする。そのためには、参加者にたくさんの解釈を生成し、選択させる過程を観察する必要がある。本研究では、発話の字義通りではない裏の意味、より具体的には意図の推測しやすさ（本研究ではこれを「理解容易度」とよぶ）を統制した実験刺激の作成が必要である。そのため、作成した各会話刺激の理解容易度について、オンラインアンケートによる評価値の分布を確認し、本実験で使用する実験刺激を選出した。

### 3.2 実験刺激作成の方針

#### 3.2.1 実験刺激の構造

実験刺激は、「状況説明文(1文)」+「2ターンの会話(発話→応答)」を1試行として構成した。「2ターン」とは、第1発話と相手の応答=第2発話を指す。参加者には第1発話の話者の立場になって、第2発話の話者の発話（第2発話）の意図を推測してもらい、解釈を回答するように指示した。

状況説明文は、推論に必要な最小限の文脈（場所、登場人物、直前の行為）を

提供する。例として、「あなたとサトシはレストランでメニューを見ています」という状況文は、場所（レストラン）、登場人物（サトシ）、行為（メニューを見る）を明示する。第1発話は、誘い・依頼・提案など、相手に何らかの応答を求める発話とした。第2発話は、字義的には第1発話に直接答えていないが、推意として何らかの意図を伝える発話とした。

### 3.2.2 統制した要素

実験刺激間の比較を可能にするために、以下の変数を統制した。

1. 語数：状況文は15~25文字以内、第1発話は15~25文字以内、第2発話は20~30文字以内に統制した。
2. 文体：すべて常体に統一した。
3. 登場人物：2名に固定し、名前はカタカナで表示した
4. 感情語：屈辱語や強い怒りの表現などがないようにした

これらの統制により、特定の実験刺激だけが解釈を歪めるリスクを低減した。

### 3.2.3 作成手順

実験刺激作成は以下の手順で行った。

#### 1. 場面の選定

Sperber and Wilson (1986/1995), および箕輪(2025)を参考に、日常的であるが真の解釈(1つの解釈には留まらない)を持たない会話例を6種類作成した。

#### 2. 初稿の作成

各場面について、状況文と2ターンの会話を作成し、計30個の会話刺激(6場面×5パターン)を作成した。5つのパターンは1→5に向けて一部の単語を変更し、解釈が難しくなるように設定した。例として、「この揚げ物頼んでみようか」という第1発話に対し、第2発話では「この揚げ物の[A]は高いですね」と返答する。[A]には、①人気、②カロリー、③雰囲気ポイント、④偏差値、⑤物語性、と解釈が難しくなるように作成した。会話刺激の作成にあたっては、刺激1はSperber & Wilson (1986/1995:194)、刺激5は箕輪(2025)を参考にし、他(2・3・4・6)は自作した。

#### 3. 改稿

初稿に対し、以下の観点から指導教員と共に改稿を行った。

- 解釈難易度の調節：推論を行わずとも理解できる難易度にならないように調整した。
  - 感情語の強度：侮辱語や極端な感情表現を削除した。
  - 話者意図の露出度：第2発話が意図を直接表現していないことを確認した。
4. 刺激の選出：理解容易度を Web アンケートにより、各会話刺激の理解容易度評定を収集し、評定値の分布を確認した上で、本実験で使用する刺激を選出した。実験刺激は理解容易度を3水準（高・中・低）に位置づけられるように分類した。詳細は3.3節で述べる。

### 3.3 実験刺激の理解容易度分布の確認

#### 3.3.1 参加者・手続き

参加者は著者・指導教員の知人および、その知人の紹介によるスノーボールサンプリングにより募集した日本語母語話者42名が回答した。5章以降で述べる本実験の参加者と重複はなかった。なお、教示効果を避けるため、参加者には研究の目的や仮説を伏せた状態で回答を求めた。実施期間は2025/11/08~10であった。

#### 3.3.2 手順

調査では、6種類の状況（散歩、クジ引き作成、勉強中、パーティーの打ち合わせ、誕生日プレゼントの選定、レストランでメニュー選び）それぞれに対して5種類の応答を組み合わせた、計30試行を提示した。各試行は、「状況説明文(1文)」+「2ターンの会話（発話→応答）」で構成した。参加者には、各項目の第2発話（相手の応答）について、「相手の意図がどの程度理解できるか」を5件法で評定するよう求めた。参加者への指示文は以下の通りである。

「以下の課題では、テキストで説明される状況と会話について評価してもらいます。【質問】この会話で、相手の返答は、あなたの発言に対して『なにかを伝える意図を持った応答』として、どの程度理解できますか？【評価基準】相手は質問に字義通りには答えていません。しかし、間接的に何かを伝えようとしています。その『相手の意図や意味』が、どのくらい容易に分かるかを教えてください。」

回答は次の選択肢から一つを選ぶよう指示した。選択肢は以下の5つである。

1. 理解できない(何を言いたいのかわからない)
2. やや理解できない(意味が分かるまでに時間がかかる)
3. 理解に迷う(理解できるとも、できないとも言えない)
4. やや理解できる(少し考えれば意味が分かる)
5. 理解できる(何を言いたいのかわ直ぐに意味が分かる)

参加者には、回答に正解・不正解はないことを明示し、「会話を読んだ直後の印象で回答すること」および「各設問は10秒以内を目安に判断すること」を求めた。設問の提示順序は、Google Formsのランダム機能を用いて回答者間でランダム化し、順序効果の影響を低減した。

### 3.3.3 分析方法

30試行の会話刺激を回答結果に応じて、実験刺激を理解容易度：高・中・低の3群に分類した。分類基準は以下の通りである。

#### ● 理解容易度(高)

対象：理解容易度の平均値  $M \geq 4.5$  かつ標準偏差  $SD < 1.0$

意図の推測が容易であり、参加者間で評定が一致している実験刺激。

#### ● 理解容易度(中)

対象： $1.5 \leq M \leq 3.2$  かつ  $SD < 1.6$

意図の推測に中程度の推論が必要であり、参加者間で評定が分かれる実験刺激。

#### ● 理解容易度(低)

対象： $M \leq 1.5$  かつ  $SD > 1.0$

意図の推測が困難であり、参加者間で「理解できない」という評定が一致している実験刺激。

これらの基準は、5件法の中央値(3.0)を基準とし、高群は「容易に理解できる」範囲 ( $M \geq 4.5$ )、低群は「理解できない」範囲 ( $M \leq 1.5$ )、中群はその中間と設定した。標準偏差 (SD) は、参加者間の評定が大きく割れていないを確認するため指標として用いた。

本実験で使用する実験刺激は、6種類の状況からそれぞれ1つずつ、計6つを選定した。選定は、(1)6種類の状況すべてから1つずつ選ぶ制約のもとで、(2)平均値と標準偏差が上記の基準に合致する刺激を候補とし、(3)最後にヒストグ

ラムの形状を確認して決定した。ヒストグラムの確認では、理解容易度を「意図が容易に理解できるか」として操作するという前提を踏まえ、回答が二群に分かれて別種の処理を行っている可能性を示す二峰性がないものを選択した。

## 3.4 結果

### 3.4.1 記述統計量

オンラインアンケートで得た30項目の会話刺激の評定値についての記述統計量を算出した（表 3.4.1）。

表 3.4.1 理解容易評定値の記述統計量

会話刺激	M	SD	MD	IQR	Min	Max
1-1	4.69	0.52	5	1	3	5
1-2	4.76	0.76	5	0	2	5
1-3	4.43	0.94	5	1	1	5
1-4	1.83	1.03	2	1	1	5
1-5	1.57	1.04	1	1	1	4
2-1	4.57	0.77	5	1	2	5
2-2	3.62	1.45	4	2	1	5
2-3	2.14	1.32	2	2	1	5
2-4	3.19	1.40	4	2	1	5
2-5	1.50	0.94	1	1	1	4
3-1	3.07	1.55	3	2.75	1	5
3-2	2.29	1.42	2	2.75	1	5
3-3	2.31	1.39	2	2.0	1	5
3-4	2.33	1.43	2	2.0	1	5
3-5	1.50	1.04	1	0	1	5
4-1	3.52	1.60	4	2	1	5
4-2	1.62	1.03	1	1	1	4
4-3	1.40	0.77	1	0.75	1	4
4-4	1.29	0.81	1	0	1	5
4-5	1.29	0.83	1	0	1	5
5-1	4.62	0.91	5	0	1	5
5-2	3.52	1.15	4	1	1	5
5-3	2.21	1.35	2	2	1	5
5-4	1.52	1.02	1	1	1	5
5-5	1.50	0.99	1	0.75	1	5
6-1	4.50	0.92	5	1	1	5
6-2	4.67	0.53	5	1	3	5
6-3	3.43	1.40	4	3	1	5
6-4	2.83	1.45	3	3	1	5
6-5	2.36	1.39	2	2	1	5

M：平均，SD：標準偏差，MD：中央値，MAX：最大値，MIN：最小値

会話刺激の数値は「種類－理解容易評定値」

### 3.4.2 実験刺激の選出

オンラインアンケートで得た理解容易度評定の詳細を表 3.4.2, 本実験で使用する実験刺激の理解容易度評定を表 3.4.3, 実験刺激の内容を表 3.4.4 にそれぞれ示す。また, 選定された各刺激の理解容易度評定分布を図 3.4.1 に示す。

3.3.3 節で設定した基準に基づき, 会話刺激から理解容易度: 高・中・低にそれぞれ 2 状況ずつ選出した。その結果, 理解容易度: 高は 6 項目, 理解容易度: 中は 18 項目, 理解容易度: 低は 6 項目であった。

表 3.4.2 理解容易評定値の判別結果

会話刺激	理解容易度
1-1	高
1-2	高
1-3	中
1-4	中
1-5	中
2-1	高
2-2	中
2-3	中
2-4	中
2-5	低
3-1	中
3-2	中
3-3	中
3-4	中
3-5	低
4-1	中
4-2	中
4-3	低
4-4	低
4-5	低
5-1	高
5-2	中
5-3	中
5-4	中
5-5	低
6-1	高
6-2	高
6-3	中
6-4	中
6-5	中

会話刺激の数値は「種類—理解容易評定値」

表 3.4.3 選定された実験刺激の理解度標定

問題番号	実験番号	理解容易度	M	SD
1-1	A	高	4.69	0.52
6-2	B	高	4.67	0.53
2-4	C	中	3.19	1.40
3-1	D	中	3.07	1.55
4-5	E	低	1.29	0.83
5-5	F	低	1.50	0.99

表 3.4.4 選定された実験刺激の詳細

刺激	理解容易度	種類	テキストの内容
A	高	状況	あなたとジョンは散歩しています
		あなた	メルセデスを運転したことある？
		相手	私は高級車を絶対に運転しないの
B	高	状況	あなたとサトシはレストランでメニューを見えています
		あなた	この揚げ物頼んでみようか
		相手	この揚げ物のカロリーは高いですね
C	中	状況	あなたはカスミとクジ引きを作っています
		あなた	ランダム化しておいた
		相手	運命に強い設計だね
D	中	状況	あなたはハルトと勉強しています
		あなた	ついでにこの課題も
		相手	机の上に載る課題の数は数えたことある？
E	低	状況	あなたはサクヤとパーティーの打ち合わせを行っています
		あなた	最初からすべて話すね
		相手	カブトムシが1匹いると道に迷わない気がする
F	低	状況	あなたとレイカは友達のお誕生日プレゼントを選んでいきます
		あなた	この花束はどうか？
		相手	この花束は言語そのものが使われているね

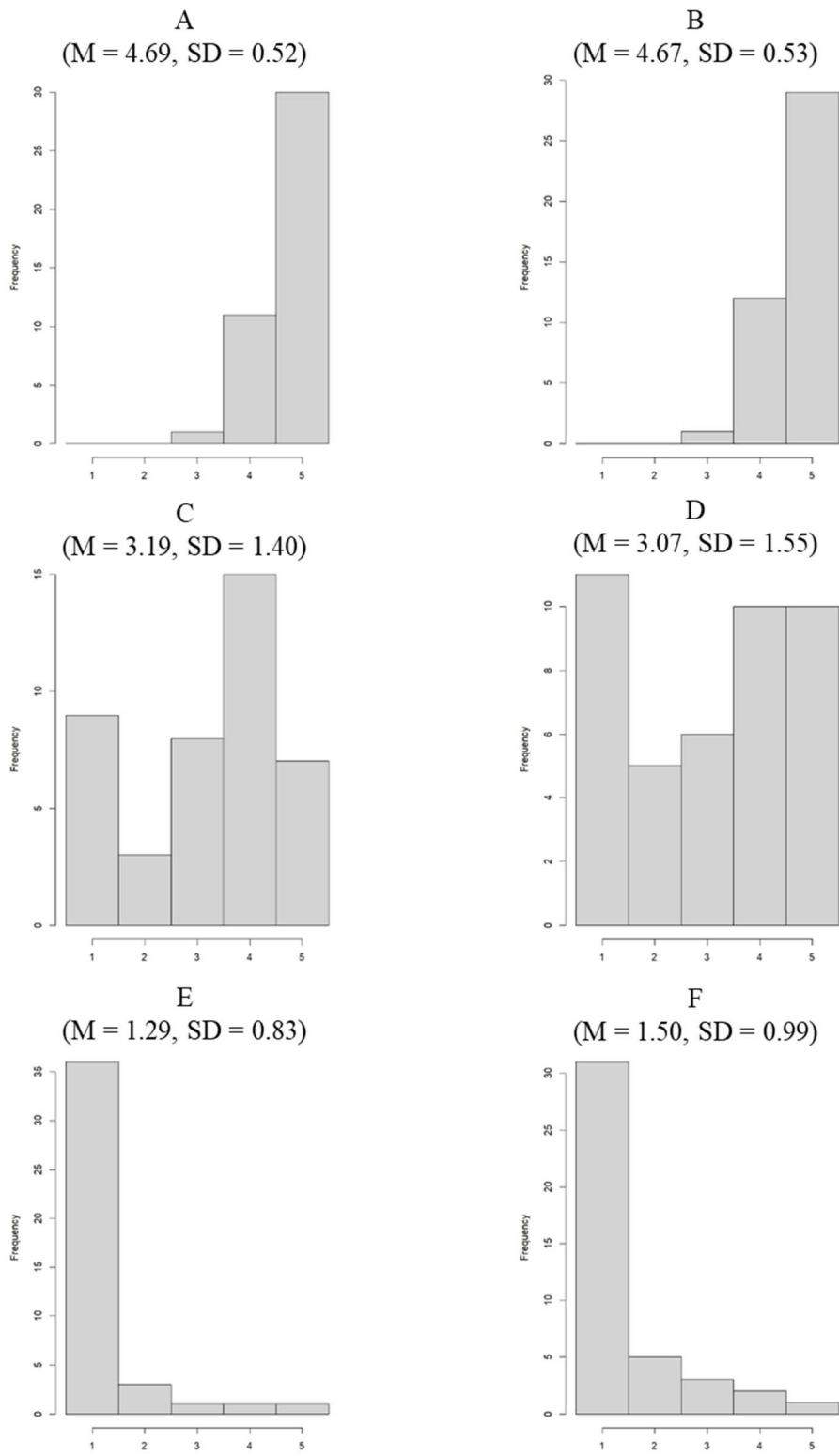


図 3.4.1 各実験刺激の理解容易度標定分布  
縦軸：回答数，横軸：理解度評定

選定された刺激は以下の通りである。理解容易度(高)の刺激(1-1, 6-2)では、評定が「4(やや理解できる)」から「5(理解できる)」に集中しており( $M \geq 4.67$ ,  $SD \leq 0.53$ )、参加者の大半が相手の意図を容易に推測できたことが推定された。理解容易度(中)の刺激(2-4, 3-1)では、評定が「2」から「4」に分散しており( $2.8 \leq M \leq 3.2$ ,  $1.40 \leq SD \leq 1.55$ )、高のように4・5へ集中して簡単過ぎるとも、低のように1・2へ集中して難しい過ぎるとも判断されていないことが推定された。理解度(低)の刺激(4-5, 5-5)では、評定が「1(理解できない)」から「2(やや理解できない)」に集中しており( $M \leq 1.50$ ,  $SD \leq 0.99$ )、参加者の大半が相手の意図を推測困難と評価したことが推定された。30種類の実験刺激内容および理解容易度標定分布をそれぞれ付録AとBに示す。

## 3.5 考察

### 3.5.1 理解容易度調査の妥当性

本調査は発話の字義通りではない裏の意味、より具体的には意図の推測しやすさ(理解容易度)を統制した実験刺激の作成するために、理解容易度評定調査を行い、結果に基づいて本実験のための刺激を選定した。結果、理解容易度(高)の刺激A・Bは $M \leq 4.60$ ,  $SD \geq 0.55$ 、理解容易度(低)の刺激E・Fは $M \geq 1.50$ ,  $SD \geq 0.98$ となり、意図した通りの統制ができていることが確認された。理解容易度(中)の刺激C・Dについては、C:( $M=3.19$ ,  $SD=1.40$ )、D:( $M=3.07$ ,  $SD=1.55$ )となり、参加者間で理解のしやすさにばらつきが見られた。これは、解釈の余地が複数存在し、参加者によって異なる推論が行われたことを示唆しており、本実験で多様な解釈を生成させるという目的に適した刺激であると考えられる。

本予備調査で作成した刺激は、Grice(1975)の協調の原則における公準違反の観点から整理できる。理解容易度(高)の刺激A・Bは、主に様態の公準(明確に言う)に違反していると考えられる。例として、刺激Bでは「この揚げ物頼んでみようか」という提案に対し、「この揚げ物のカロリーは高いですね」とYes/Noで直接答えず間接的な表現を用いている。しかし、発話内容と文脈の関連性は保たれているため、推論経路が限定されることが考えられる。一方、理解容易度(中・低)の刺激C~Fは、主に関係の公準(関連性のあることを言う)に違反していると考えられる。特に刺激E・Fでは、第1発話と第2発話の関連性が極めて不明瞭であり、明示的には関連性が無いが何かしらの意図があると解釈を多く

生成すると考えられる。

### 3.5.2 本実験への接続

以上の結果と考察を踏まえ、選定された 6 刺激を発話解釈に関する実験で使用する。しかし、オンラインアンケートによる理解容易評定と、本実験における認知的負荷が対応しているかは未確認である。そこで本研究では、理解容易度の代理変数として操作が機能しているかを確認するため、解釈生成までの反応時間を測定する。理解容易度が高い刺激では解釈が容易に生成できるため反応時間は短く、理解容易度が低い刺激では解釈を探索的に推論する必要があるため反応時間は長くなると予測される。この反応時間の差異を確認することで、理解容易度の操作チェックとする。

# 第4章 推論形式判別基準作成のための予備実験

第3章では、理解容易度を制御した本実験で使用するため会話刺激を作成し、Web アンケートによる評定により、理解容易度の3群に沿った実験刺激を計6つ選出した。本章では、予備実験を実施して Think Aloud 法によって得られた参加者の回答から、推論形式を判別するための基準について述べる。

## 4.1 目的

本予備実験では、参加者から得られた Think Aloud の回答から推論形式を分類できる基準を作成することを目的とした。同時に、Think Aloud 法によって、参加者の推論プロセスが十分に可視化できるかを確認する。

ここでは分類対象の推論形式として、通常の帰納・演繹・アブダクションに加え、認知的共感・情動的共感を含めた。共感は通常は「推論」に含まれないが、2章で見たように、会話における発話理解には重要な役割を果たしているため、この方略も用いていることが想定されるからである。

## 4.2 方法

### 4.2.1 参加者・手続き

参加者は、北陸先端科学技術大学院大学石川キャンパスの日本語母語話者2名（男性2名、 $M=24.5$ ,  $SD=0.50$ ）であった。この2名は5章以降で報告する本実験には参加していない。なお、参加者には研究の目的や仮説を伏せた状態で回答を求めた。実施期間は2025/05/24-2025/05/27であった。

### 4.2.2 実験刺激

実験刺激は、第3章で作成した会話刺激とは異なる、3つの会話刺激を使用した。これらの刺激は、水野(2024)を参考に作成した。各刺激は、「状況説明文」+「2ターンの会話」+「追加情報①②」で構成した。構成の詳細は以下の通りである。

1. 会話文: 状況説明文と2ターンの会話(実験刺激と同様の構造)
2. 追加情報①: 会話の背景を補足する情報(例: 登場人物の状態, 場所の詳細)
3. 追加情報②: 第2ターン発話の意図を推測する手がかり

追加情報を段階的に提示した理由は、参加者の持つ情報が少ない状態での推論と、情報が増えた後の推論を、それぞれ観察するためである。これにより、情報量が推論形式に与える影響を検討できる。また、本試行の前に、1つの練習試行を実施し、Think Aloud 法の回答方法に慣れさせた。

#### 4.2.3 手法

実験は参加者ごとに個別に実施した。参加者は、Google Forms で提示される刺激を読み、Think Aloud 法で推論プロセスを言語化した。実験の流れは以下の通りである。

##### (1) 会話文の提示

参加者に会話文を提示し、「この会話で、2人目がなぜこのような発話を行ったかを考えてください」と教示した。

##### (2) Think Aloud 法の実施

参加者には、以下の教示を与えた。

「思いつく限り、2人目の発話にどのような意図があるかを、声に出して答えてください。沈黙が10秒続いた場合、実験者から『他に思いつきますか?』と声をかけます。この問いかけには返答しなくて結構です。回答の最初には『1つ目は～』と意図の数が分かるように発言してください。回答が思いつかない場合は、『以上です』と発言してください。」

##### (3) 追加情報①の提示

参加者が「以上です」と発言した後、追加情報①を提示し、再度(2)と同じ手続きを実施した。

##### (4) 追加情報②の提示

追加情報②を提示し、再度(2)と同じ手続きを実施した。

##### (5) 半構造化インタビュー

1つの会話刺激が終了した後、これまでに述べた解釈の中から「最も妥当と思われる意図を1つ選び、その理由を説明してください」と求めた。実験は、練習で(5)を除く1回と、本番で(1)～(5)を2回実施した。データはICレコーダーで音声録音し、後日Microsoft Wordのトランスクリプト機能を用いて

文字起こしを行った。

#### 4.2.4 判別基準の作成方法

文字起こしされた回答を，以下の手順で推論形式として分類した。

##### (1) 分析単位の設定

1つの解釈ごとに分析を実施した。1つの解釈の範囲は，実験指示で与えた「1つ目は」という発言から次の「2つ目は」の前までとした。その際，「えっと，あー」などフィラーは除いた。

##### (2) 推論形式の分類

理論的定義(第2章)に基づき，5カテゴリーを作成した（表4.2.1）。

表 4.2.1 推論形式の判別基準

分類	基準
事例 (帰納に対応)	経験語（参加者の経験に基づく一般化）or 事例提示(参加者の経験談) or 観察（頻度・数量）
規則 (演繹に対応)	規則語(一般化された規則[ルール・定義]) or 規範的判断（行為に対する規範・ルール）
推定 (アブダクション に対応)	仮定(もし/だとすると) or 反実仮想（～していたら…なはずだ）+可能性語（～かも） 候補列挙（A か B か） / 仮定+説明志向語(～と説明できる) 状況や会話テキストには無い設定
認知的 共感	視点取得句（相手の立場なら/彼の視点では）+心的状態動詞(～と思う) シミュレーション句（自分なら～→彼も）+根拠提示
情動的 共感	情動語（つらい/悲しい/嬉しい）+共鳴表現（共感する身体手掛かり） +情動推定（～そうだ）

分析は筆者と指導教員とともに2名の参加者の回答を読み，上記の5カテゴ

リーに分類した。両者で分類が一致しなかった箇所については、協議の上で判定した。

## 4.3 結果

### 4.3.1 回答の言語化

2名の参加者は、本番で回答した2つの会話刺激に対して、合計で課題1：15個、課題2：14個の解釈を生成した(1刺激あたり平均14.5個)。追加情報による解釈生成数の変化も含め、回答数を表4.3.1に示す。追加情報の有無で生成数は増加すると予想していたが、解釈の生成数は現象していった。

Think Aloudによる発話では、沈黙が10秒続く場面はほとんどなく、参加者の1人は自発的に推論プロセスを言語化していた。2人目は沈黙が5秒以上起きることは無く「ん～」などの発言により考えていることは示されたが、回答数は低かった。

表 4.3.1 追加情報による解釈候補生成数の変化

課題	課題1	課題2	課題1	課題2	計
参加者	A	A	B	B	A+B
追加なし	5	5	1	1	12
追加情報①	4	4	2	1	11
追加情報①②	2	2	1	1	6
計	11	11	4	3	29

### 4.3.2 推論形式の判別結果

推論形式の出現頻度を表4.3.2示す。結果より、以下の2点が明らかになった。

1. 複数の推論形式が組み合わせられる。当初、本研究は「発話解釈は単一の推論形式で行われる」と想定していたが、実際には、複数の推論形式が組み合わせられることが頻繁に観察された。特に、「推定+認知的共感」(アブダクション+認知的共感)が課題1で10回、課題2で7回出現し、これが最も頻繁な組み合わせであった。これは、発話解釈が単一の推論形式ではなく、複数の推論形式を統合的に用いるプロセスであることを示唆する。
2. 演繹も一定頻度で出現している。「規則」(演繹)が単独で課題2に3回出現し、「認知的共感」が単独では出現しなかったものの、「推定+共感」「推定+規則+共感」といった組み合わせで一定頻度出現した。これは、演繹と

認知的共感が、アブダクションを補完する役割を果たしていることを示唆する。

しかし、次の問題が明らかになった。一つ目は、「事例」と「規則」の区別が困難であることである。二つ目は、「情動的共感」の出現頻度が極めて低いことである。「情動的共感」に該当する発話は、2名の参加者の回答全体でわずか2回しか出現しなかった。

表 4.3.2 各推論形式の出現頻度

課題	課題 1	課題 2	課題 1	課題 2	計
参加者	A	A	B	B	A+B
推定	4	3	0	0	7
規則	0	2	0	0	2
事例	0	0	0	0	0
認知的共感	0	2	0	0	2
推定+認知的共感	5	4	4	3	16
推定+規則+認知的共感	1	0	0	0	1
推定+認知的共感+情動的共感	1	0	0	0	1
計	11	11	4	3	29

## 4.4 考察

### 4.4.1 回答の言語化

予測していた「追加情報が増えることで解釈候補の生成数は増加する」は支持されなかった。この結果の要因は2つ考えられる。一つ目は、関連性理論の「十分な認知効果を、過度な処理努力なしに得られる」ことを踏まえると共通の情報が得られているため少ない労力で解釈を得ようとするため追加情報を得ることで生成数も減少すると考える。二つ目は、追加情報がない場合には共通の情報を補おうと探索的に解釈候補を生成していくが、追加情報を得ることで探索範囲が狭くなり、生成する必要がなくなったのではないかと考える。以上を踏まえて発話解釈に関する実験では追加情報を提示しないこととした。

### 4.4.2 判別基準の妥当性

「事例（帰納）」と「規則（演繹）」の区別が極めて困難であった。例えば、看板を見ながら「この先、動物注意だって」という発話は、以下のどちらとも解釈

できる。

- 事例（帰納）：「これまで『動物注意』の標識がある場所では、実際に動物が出た場面を何度も見た。だから、ここでも動物が出そうだ。」（事例の蓄積→一般化→適用）
- 規則（演繹）：『動物注意』という標識は、動物が出没する可能性がある場所に設置されるというルール（制度的な取り決め）がある。だから、ここは動物が出る可能性がある。」（一般ルール→適用）

この2つはどちらも、何かしらのルール（知識）に基づいた推論である。すなわち、「経験→ルール→適用」という推論なのか、「制度的・自然的・論理的な一般ルール→適用」という推論である。このルールが、参加者の経験に基づくものなのか一般的なルールなのかは、参加者にインタビューで確認するなどしなくては判定不能だと思われる。

この問題は、理論的にも興味深いと考える。人間の推論プロセスは、論理的に厳密な形式(帰納 vs 演繹)ではなく、経験とルールが混在した形で機能している可能性がある。そこで本研究では、事例と規則を「規則」として包括的に扱うこととする。

また本予備実験から得た知見は、発話解釈において複数の推論形式が組み合わされるといふ点である。当初、本研究は「発話解釈は主にアブダクションで行われる」という仮説を持っていたが、実際には、アブダクションだけでなく、演繹や認知的共感も同時に用いられることが明らかになった。本実験での分析ではこの結果を活かした分析を実施する。

#### 4.4.3 発話解釈に関する実験への接続

本予備実験の結果を踏まえ、本分析では次の変更を行う。

##### 1. 判別基準の見直し

本実験では判別基準を表 4.4.1 のように更新して使用した。

##### 2. Think Aloud 手続きの改善

本予備実験では沈黙が 10 秒続いた場合に実験者が「他に思いつきますか?」と声をかけたが、この実験者の声が録音に混入し、文字起こしの妨げとなった。そこで、発話解釈に関する本実験では、沈黙が 3 秒続いた場合にベルを鳴らすこととした。これにより、実験者の介入を最小限に抑え、録音の質を向上させる。

### 3. 回答時間の制限

本予備実験では回答時間に制限を設けなかったが、参加者によっては長時間考え込む場面があった。そこで本実験では、研究室内で実施した予行練習の結果を踏まえ、1刺激あたりの回答時間を3分に制限した。この時間は、十分な数の仮説が生成され、参加者の負担が過度にならず、無理にひねり出すような状況にならない、という3つを満たす最適な時間として設定した。

表 4.4.1 本分析に使用する推論形式判別基準

分類	基準
事例 (演繹)	経験語 (参加者の経験に基づく一般化) or 事例提示 (参加者の経験談) or 観察 (頻度・数量) 規則語(一般化された規則[ルール・定義]) or 規範的判断 (行為に対する規範・ルール)
推定 (アブダクション)	仮定(もしだとすると) or 反実仮想 (~していたら…なはずだ) +可能性語 (~かも) 候補列挙 (A か B か) / 仮定+説明志向語(~と説明できる) 状況や会話テキストには無い設定
認知的共感	視点取得句 (相手の立場なら/彼の視点では) +心的状態動詞(~と思う) シミュレーション句 (自分なら~) +根拠提示

## 第5章 発話解釈に関する実験

第3章では、理解容易度を高・中・低の3群に操作した刺激（会話テキスト）を作成した。第4章では、Think Aloud法によって得られた発話プロトコルから、ルールに基づいた演繹的推論、仮説を生成するアブダクション、あるいは、認知的共感を判別する基準を確立した。本章では、これらの成果を踏まえ、本実験の方法を述べる。

本章の構成は以下の通りである。5.1節では本実験の参加者の詳細を述べる。5.2節では実験デザインについて整理し、実験の仮説を提示する。5.3節では実験刺激と提示順序を説明する。5.4節では実験手続きについて述べる。5.5節では直感回答課題の詳細と質的分析の方法を示す。5.6節では多数回答課題の詳細と質的分析の方法を示す。5.7節では解釈選択課題の詳細と質的分析の方法を示す。5.8節では個人差測定の詳細を示す。5.9節では半構造化インタビューの詳細を述べる。5.10節では本実験で得られた結果の一般化可能性を評価するための統計手法について述べる。

本実験は、北陸先端科学技術大学院大学ライフサイエンス委員会の承認(承認コード: 人07-026, 承認日: 令和7年10月7日)を得て、「人を対象とした研究の実施に関する規則」に従い実施した。

### 5.1 参加者についての詳細

#### 5.1.1 募集方法

本実験は北陸先端科学技術大学院大学「人を対象とした研究の実施に関する規則」に従い、参加者の自由意志に基づく自発的参加の同意を得てから実施した。参加者は北陸先端科学技術大学院大学石川キャンパスの学生（博士前期課程・後期課程・研究生）を対象としたメーリングリストを通じて募集した。参加は任意であり、参加しないことによる不利益はないことを明示した。参加条件は日本語母語話者とした。謝金は北陸先端科学技術大学院大学の規定に従い、1時あたり1,300円を支払った。また、Think Aloudの発話による負荷を軽減するため、参加者には500mLのミネラルウォーターとのお菓子を2つ配布した。ミネラルウォーターは、実験の休憩中に飲むように指示し、実験終了後にのお菓子を配布した。

#### 5.1.2 参加者と実施期間

参加者は、北陸先端科学技術大学院大学石川キャンパスの学生で、日本語を母

語とする者 43 名であった。そのうち、最終的な分析対象は 38 名（男性 28 名（ $M = 24.3$ ,  $SD = 1.11$ ）、女性 10 名（ $M = 24.7$ ,  $SD = 1.89$ ））であった、実験手続きの不備により 3 名、実験途中の申し出による中止により 2 名の、合計 5 名のデータを分析から除外した。実施期間は 2025/11/27～2025/12/20 であった。

## 5.2 実験デザイン

### 5.2.1 目的

本実験の目的は、発話解釈がどのような推論形式で行われているのかについて、理解容易度に応じて推論形式の使用パターンがどのように変化するかを明らかにすることである。また、心理測定尺度を用いて、推論形式の選択に関する個人差を検証する。

### 5.2.2 仮説

本実験では、以下の仮説を検証する。

仮説 1：理解容易度が低いほど、アブダクション推論の使用頻度が増加する

仮説 2：コミュニケーションと思考に関する個人特性は、推論形式の使用頻度と関連する

### 5.2.3 実験計画

本実験は、1 要因（理解容易度：高・中・低）参加者内デザインである。各参加者は、理解容易度高・中・低それぞれ 2 刺激合計 6 つの刺激すべてに対して、直感回答課題、多数回答課題、解釈選択課題を順に実施した。課題については 5.5～5.7 節で説明する。

以下の変数を設定した。

- 従属変数：各推論形式（帰納・演繹、アブダクション、認知的共感）およびその組み合わせの頻度・発話解釈にかかる時間
- 個人特性：心理測定尺度（心の理論・コミュニケーション・スキル）
- 思考特性：心理測定尺度（認知的欲求・認知的完結欲求）
- 性格特性：心理測定尺度（短縮版 Big Five）

## 5.3 実験刺激

### 5.3.1 実験刺激の内容

実験刺激は、第 3 章で選定した 6 つの会話刺激を使用した（表 5.3.1 に再掲）。

これらの刺激は、理解度(高・中・低)の各群から 2 刺激ずつ選定されており、各刺激は「状況説明文」+「2 ターンの会話(発話→応答)」で構成されている。

表 5.3.1 実験刺激の詳細

刺激	理解容易度	種類	テキストの内容
A	高	状況	あなたとジョンは散歩しています
		あなた	メルセデスを運転したことある？
		相手	私は高級車を絶対に運転しないの
B	高	状況	あなたとサトシはレストランでメニューを見えています
		あなた	この揚げ物頼んでみようか
		相手	この揚げ物のカロリーは高いですね
C	中	状況	あなたはカスミとクジ引きを作っています
		あなた	ランダム化しておいた
		相手	運命に強い設計だね
D	中	状況	あなたはハルトと勉強しています
		あなた	ついでにこの課題も
		相手	机の上に載る課題の数は数えたことある？
E	低	状況	あなたはサクヤとパーティーの打ち合わせを行っています
		あなた	最初からすべて話すね
		相手	カブトムシが 1 匹いると道に迷わない気がする
F	低	状況	あなたとレイカは友達誕生日プレゼントを選んでいきます
		あなた	この花束はどうか？
		相手	この花束は言語そのものが使われているね

### 5.3.2 刺激の提示順序

刺激の提示順序は参加者間でランダム化した。具体的には、6 刺激(A~F)の全順列(6! = 720 通り)から、各刺激が各提示位置(第 1~第 6 番目)に均等に出現するように、45 通りの順序を抽出した。この制約付きランダム化により、各刺激は全体として 45 回ずつ提示され、各提示位置における各刺激の出現回数は 7 回または 8 回のいずれかとなるように調整した。これにより、特定の刺激が特定の位置に偏って提示されることを防いだ。作成した提示順序表を付録 C に示す。

### 5.3.3 刺激の提示方法

刺激は PowerPoint で作成された動画として提示した(詳細は付録 D~F 参照). 各刺激の動画は, 表 5.3.1 の順序で情報を提示した. 実験課題の動画は, 開始・終了・休憩の 3 つ動画と, 各刺激の動画 (計 6 つ) を, Microsoft Clipchamp で結合して作成した. 参加者ごとに提示順序が異なるため, 45 名分の動画(45 通りの提示順序)を作成した.

表 5.3.2 実験刺激の提示と課題の順

提示順	提示文
状況	あなたとハナコは 図書館で勉強しています
1 人目の発話	あなた: ご飯食べにいかない?
2 人目の発話の予測	ハナコの返答を考えてください
2 人目の発話	ハナコ: もう, 聞かないでよ
直感回答課題	ハナコの発言の意図は何ですか? すぐにひと言で回答してください
状況~2 人目の発話	同様のテキストを提示
熟慮課題	マユミの発言の意図は何ですか? できるだけたくさん答えてください
状況~2 人目の発話	同様のテキストを提示
解釈選択課題	マユミの意図の解釈を一つずつ声に出して比べ 一番適切だと感じるものを選んでください

表 5.3.2 の例は, 課題説明で使用した刺激

## 5.4 手順

### 5.4.1 実験の流れ

実験は 1 日で実施し, 以下のスケジュールで進行した.

1. 実験の説明, 同意書の記入(約 10 分)
2. 課題の説明, 練習試行(約 30 分)
3. 本試行: 6 刺激×3 課題(約 60 分, 各刺激終了後に 2 分休憩)
4. 休憩(5 分)
5. 個人差測定(約 10 分)

## 6. 半構造化インタビュー(約5分)

### 5.4.2 インフォームドコンセント

実験開始前に参加者へ研究説明書を提示し、以下を説明した。(1) 研究の目的、(2) 課題内容と所要時間、(3) Think Aloud を含む発話の記録方法(録画の有無)、(4) 予想される負担と中止・休憩が可能であること、(5) 参加は任意であり、いつでも同意を撤回でき、不利益が生じないこと、(6) データの匿名化・保管方法と公表時の取り扱い、(7) 問い合わせ先。説明後、同意書への署名をもって参加同意を得た。説明書・同意書・同意撤回書は付録 H に示す。

### 5.4.3 実験環境

実験は、北陸先端科学技術大学院大学知識科学系研究棟Ⅱ棟 8 階(K-83) (以降、実験室)にて個別に実施した。実験室内にパーティションを用いて参加者のみが入る空間を作成し、参加者が課題に集中できる環境を整えた。実験室の環境を図 5.4.1、用いた実験機器とそれぞれの役割について表 5.4.1 にそれぞれ示す(※ぬいぐるみは参加者を想定)。実験中にこの部屋に人が立ち入ること、大きな物音が参加者に聞こえるなどの、集中を妨げる事態はなかった。

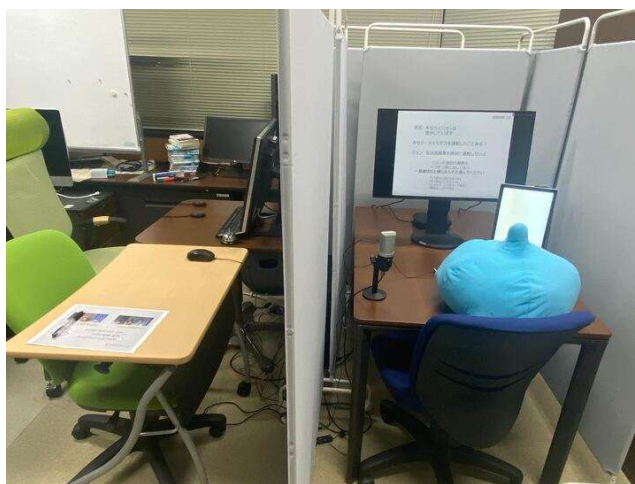


図 5.4.1 実験風景

左側が実験実施者のスペース、右側が参加者のスペース

表 5.4.1 実験機器の詳細・配置および役割

機器	配置	用途
デスクトップ PC	実験実施者側	実験操作および Zoom でのデータ収集
正面モニター	参加者正面	刺激提示用
Surface	参加者側机上	解釈選択課題時に多数回答課題の回答内容を提示
ピンマイク	参加者が装着	参加者の発話録音用 (Zoom で録画)
机上マイク	参加者側机上	解釈選択課題提示のため多数回答課題の音声入力用
IC レコーダー	参加者側机上	バックアップ録音用

実験実施者は、左側奥の椅子に着席しデスクトップ PC を操作した。実施者の介入は、休憩の確認および発話の促しなど必要最小限に限定した。

参加者の発話は Zoom で録画し、画面共有機能を用いて刺激提示画面も同時に記録した。参加者のカメラはオフにし、画面録画のみを行うことで参加者個人が特定されないよう配慮した。

多数回答課題における参加者の発話は、Windows 標準搭載の音声入力機能を用いてリアルタイムで文字起こしを行った。文字起こしの内容は Google ドキュメントを介して参加者側の Surface に共有し、解釈選択課題で参加者が自身の回答内容を確認できるようにした。

バックアップとして IC レコーダーでも録音を行った。すべてのデータはセキュリティが確保されたクラウドストレージに保存した。

## 5.5 直感回答課題

### 5.5.1 目的

直感回答課題の目的は、会話刺激に対する初期的な発話意図解釈を、時間制約下(15 秒)で引き出し、発話を解釈するための処理時間と初期解釈の内容を取得することである。時間制約を設けることで、参加者が長い検討や説明を行う前の、比較的自動的・即時的な解釈を得ることができると考えた。さらに、この初期解釈は、後続の多数回答課題および解釈選択課題で得られる解釈と比較する基準点として位置づけられる。

### 5.5.2 手順

直感回答課題では、状況説明文と 2 ターンの会話が提示された後、1 ターン目の発話を参加者が行っている (ロールプレイング) と想定し「相手がなぜこのよ

うな発言をしたか」を，制限時間 15 秒以内に，ひと言で口頭回答するよう求めた．参加者には，以下の教示を与えた．

「相手の発言の意図は何ですか？ すぐにひと言で回答してください．制限時間は 15 秒です．」

課題開始は，動画で流れるビープ音と教示の提示を合図とした．課題終了は，ベルの音を合図とした．参加者が 15 秒以内に回答しなかった場合，その試行は無効とした．

練習試行は，フィードバック付きの試行と通し練習を各 1 回実施した．フィードバック付きの試行では，参加者が回答した後，実験者が①回答方法を理解しているかの確認，②回答内容の確認(例：ひと言で回答できていない，第三者視点での回答をしてしまってる)を，補足で確認・説明した．本試行では，6 刺激に対して各 1 回，計 6 回実施した．提示したスライドを付録 D に示す．

### 5.5.3 収集するデータ

口頭による録音および回答時間を収集した．

### 5.5.4 発話解釈にかかる時間の算出

発話解釈にかかる時間は，ELAN(version 7.0)を用いて算出した．ELAN は，音声や動画データに詳細なアノテーションを多層的に付与し，分析するためのソフトウェアである．

発話解釈にかかる時間は，被験者が聞き手の発言内容からどのような意図で発言したのかを理解し発話を開始するまでの時間である．以下の始点と終点の時間間隔を用いて算出した (図 5.5.1)．

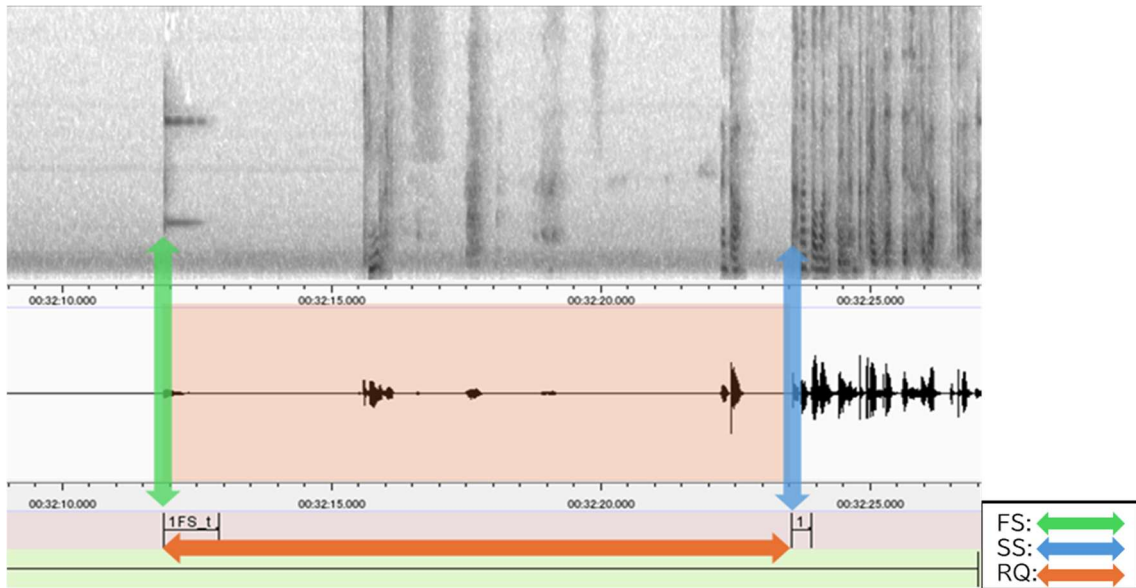


図 5.5.1 反応時間(RQ)の算出方法

1. FS：開始のピープ音
2. SS：発言の意図の1音目の波形の始点

発話解釈にかかる時間は以下の式を用いて算出した。

$$\text{発話解釈にかかる時間 [ms]} = \text{SS[ms]} - \text{FS[ms]}$$

この指標は、値が 15,000[ms]（制限時間）に近いほど、解釈に時間が掛かったことを示す。

#### 5.5.5 発話解釈にかかる時間の算出対象

発話解釈にかかる時間の算出には、以下の条件を満たす回答を使用した。

- 制限時間 15 秒以内に回答できているもの

**SSの例**

例：発言の意図は、ハナコは行きたくないと考  
えている

○ 「行きたくない」のiの音

✗ 「発言の意図は～、・・・」のhの音

図 5.5.2 発話解釈開始の判別基準

例として、「発言の意図は、ハナコは行きたくないと考えている」という回答の場合、「行きたくない」の「i」の発生波形を終点(SS)とした(図 5.5.2)。「発言の意図は」という前置きは、反応時間に含まれる。

発話解釈にかかる時間の算出には、被験者 1 人から最大 6 個対象となった。無回答や 15 秒を超えてからの回答もあり、最小 2 個であった。

### 5.5.6 発話解釈の推論形式の導出

直感回答課題で得られた初期解釈の推論形式は、以下の手順で導出した。

- (1) 多数回答課題で同じ内容の解釈が生成されているかを確認
- (2) 同じ内容の解釈が多数回答課題にある場合、その解釈の推論形式を、直感回答課題の推論形式として割り当てる、(推論形式の判別については 4.4.3 節参照)
- (3) 同じ内容の解釈が多数回答課題にない場合、推論形式を「NA(無回答)」として処理する。

この方法を採用した理由は、直感回答課題では「ひと言」での回答を求めため、推論プロセスが言語化されず、推論形式を直接判別できないためである。一方、多数回答課題では、回答時間が 3 分と長いができるだけたくさん回答するよう指示しているため、熟考できるわけでもない。したがって、直感回答と同じ推論が行われたという近似的な前提を置いている。

## 5.6 多数回答課題

### 5.6.1 目的

多数回答課題の目的は、同一の会話刺激に対して参加者が構成しうる複数の発話意図解釈(解釈仮説)をできる限り多く生成させ、その生成過程を Think Aloud 法により言語化させることで、(a)解釈仮説の生成数、(b)各仮説が依拠する推論形式、(c)生成の順序や根拠づけに現れる手がかり(状況、関係性、一般知識等)を収集することである。制限時間(3分)を設け、沈黙時の促し等を含む指示により、途中で確信が揺らぐ場合も含めて探索的に生成を継続させることで、解釈候補の生成過程を観察可能にする。

## 5.6.2 手順

多数回答課題では、参加者に対して、状況説明文と2ターンの会話から「相手がなぜこのような発言をしたか」を、思いつく限り多く生成し、その際に生じる思考過程を逐次、口頭で報告するよう求めた。回答はThink Aloud法を用いて実施した。参加者には、以下の教示を与えた。

「あなたは、ここで相手は何を伝えたいのかを、思いつく限り答えてください。回答は、思っていること・考えていることを、全て口に出し続けてください。例として、『んん～、いまハナコの意図を考えています』『いまハナコとは喧嘩していると仮定すると、話しかけられないで欲しいのかな』のように、あなたの視点から相手の発言の意図を考えてください。」

課題遂行時の注意点として、以下の3点を明示した。

1. 解釈が「違うかもしれない」と感じても、発話を止めずに考え続けること。  
探索的な仮説生成を促進するためである
2. 途中で「違う」と判断した場合は、その場で言い直しをしてよいこと。  
参加者が「最初の言い方が適切でなかった」と感じた場合でも、「先ほどの解釈は違って、本当はこう言いたかった」という形で修正しながら発話を継続してよいことを明確化した。発話の萎縮や沈黙を抑制するためである。
3. 沈黙が3秒以上続いた場合、ベル音を鳴らすため、その時点で何を考えていたかを説明すること。  
予備実験の結果を踏まえ、実験者の介入を最小限にするため、自動ベル音を採用した。また、本課題には正解がないことを伝え、参加者が自由に発話できるよう「あなたがどう考えたかを知りたいので、思いついたことをどんどん口に出してください」と付言した。

回答方法として、生成した解釈の開始時点と個数が判別できるように、「1つ目は～、2つ目は～」の形式で列挙するよう求めた。制限時間は3分とし、経過時間の目安として1分ごとにチャイム音(「チン」)を提示した。

練習試行では、フィードバック付きの試行と通し練習を各1回実施した。本試行では、6刺激に対して各1回、計6回実施した。提示したスライドを付録Eに示す。

### 5.6.3 収集するデータ

口頭発話を録音し，その内容の文字起こしを行った。

### 5.6.4 発話解釈の推論形式の導出

多数回答課題の発話プロトコルから推論形式を導出する手順は，以下の通りである。

手順 1: 音声録音を文字起こしし，発話プロトコルを作成した

手順 2: 第 4 章の表 4.4.1 の判別基準に基づき，各解釈に推論形式のラベルを付与した

手順 3: ラベル付けは，筆者と指導教員が協議の上で実施した

ラベル付けの例を図 5.6.2, 5.6.3, 5.6.4 に示す。

### 5.6.5 発話解釈の推論形式の分析対象

発話解釈の推論形式の同定には，以下の条件を満たす回答のみを使用した。

- 推論形式が判別可能な発話解釈の意図
- 同じ内容の解釈は最初に出現したもののみを対象とする(重複を除く)

発話解釈の推論形式の導出には，最大 9 個，最小 1 個が対象となった。

多数回答課題の 1 刺激における推論の分析対象について述べる。1 つの解釈は図 5.6.1 のような構造をしていると考えられる。2 ターン目の発話(図中の「発言」)に対する解釈を説明可能にするための推論が行われていると考えており，その部分(図中のオレンジで囲んだ部分)が本研究における推論形式の分析対象である。各推論形式のラベル付け例を図 5.5.4 (推定)，図 5.5.5 (規則)，図 5.5.6 (認知的共感) に示す。

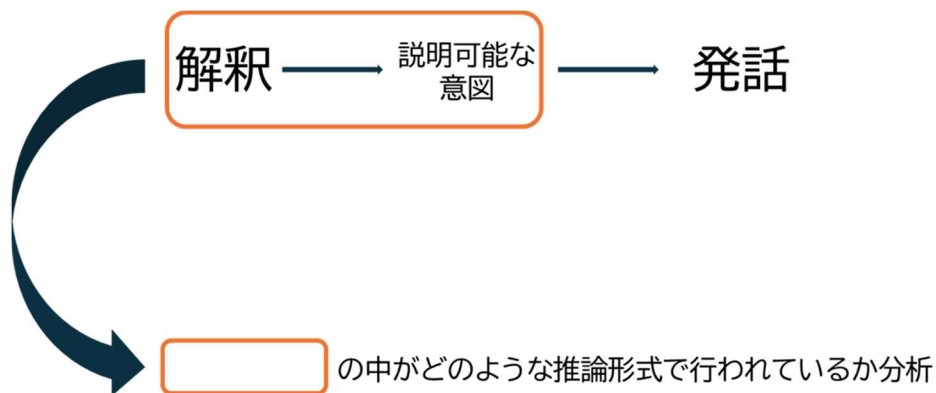


図 5.6.1 回答分析対象

参加者の回答

アブダクション

カロリーを比較したいという主張で、  
一旦カロリーに着目しませんか？というふうに言いたい

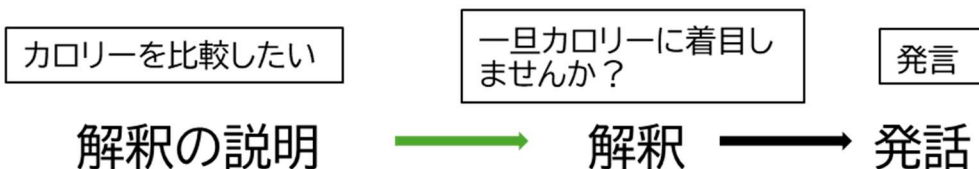


図 5.6.2 回答の分析（推定の場合）

参加者の回答

演繹

四つ目として、なんか逆にサトシが減量とか増量してる側じゃなくて、  
私がいつもダイエットとかしてる側でサトシがこの揚げ物もカロリー高  
いじゃんっていうことで、ダイエットしてないのに、こんなの食べて大丈夫  
なの？みたいな私のためにをやめた方がいいんじゃない？って進め  
てる意図

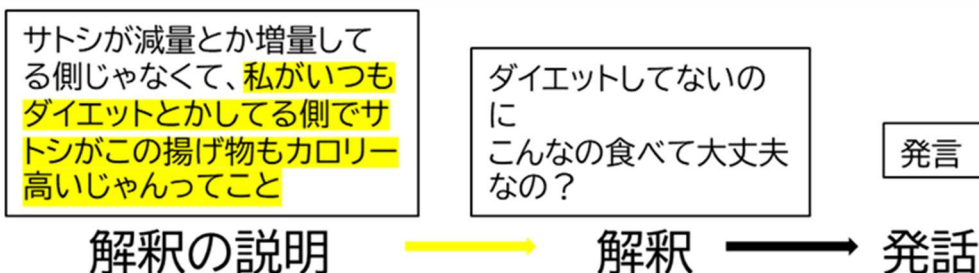


図 5.6.3 回答の分析（規則の場合）

参加者の回答

認知的共感

四つ目私が枕言葉でつけて言っているからハルトとしては、たくさん  
行ってるから、めんどくさいなっていう意味

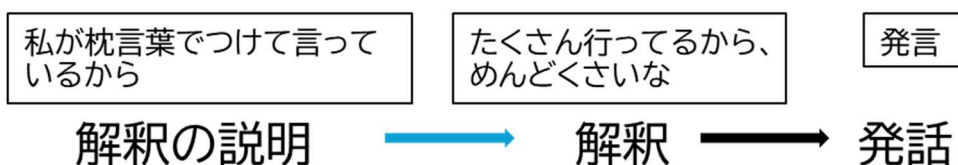


図 5.6.4 回答の分析（認知的共感の場合）

## 5.7 解釈選択課題

### 5.7.1 目的

解釈選択課題の目的は、多数回答課題で参加者自身が生成した複数の解釈仮説を対象として、それらを比較・評価した上で最も適切だと判断する解釈を1つ選択させ、選択に至る評価基準および正当化を言語化させることである。本課題を設けることで、解釈仮説生成(直感・多数回答課題)と区別された仮説選択段階を明確化し、(a)最終的に採択される解釈の内容と推論形式、(b)生成された候補の中で何が棄却され、何が残るか、(c)「適切さ」を支える根拠(状況整合性、対人配慮、一般的期待など)を収集できる。さらに、直感回答課題で得た初期解釈との一致・不一致を比較することで、時間制約下の初期解釈が最終選択にどの程度維持・修正されるかを検討可能にする。

### 5.7.2 手順

解釈選択課題では、多数回答課題で参加者が生成した複数の解釈候補の中から、最も適切だと考える解釈を1つ選択し、選択した理由を口頭で回答するよう求めた。参加者には、以下の教示を与えた。

「設問2(多数回答課題)で考えた相手の意図を、一つずつ声に出して比べ、解釈として一番適切だと感じるものを選んでください。」

多数回答課題で生成された解釈候補は、課題提示用モニターの右下に設置した補助端末(Surface)の画面に提示した。具体的には、多数回答課題の口頭回答を別端末(実験実施者側端末)上にリアルタイムで文字に起こし、Google ドキュメントとして整形した上で、解釈選択課題開始時に補助端末画面に提示した。これにより、参加者が多数回答課題で挙げた候補を見返しながら、比較・選択を行えるようにした。この方法を採用した理由は、候補数が多い場合に、参加者が複数の解釈候補を保持しつつ比較・評価するための作業記憶(ワーキングメモリ)への負荷が過度に高まり、解釈の選択が「記憶のしやすさ」や想起の偶然性に左右されることを避けるためである。また、表示領域の制約により一部の候補が画面外に隠れる場合があるため、参加者が指でスクロールして閲覧することをあらかじめ許可した。

回答形式は、候補を1つずつ取り上げて適否の理由を述べた上で、最終的な選択と根拠を明示する形とした。教示では、例として以下のように説明した。

「1つ目に出した〇〇は、××だから合っている。2つ目に出した〇〇は、△△だから違う。1つ目と2つ目なら、1つ目のほうが合っている。だから、1つ目の〇〇である。理由は××だからだ。」

制限時間は3分とした。練習試行では、フィードバック付きの試行と通し練習を各1回実施した。本試行では、6刺激に対して各1回、計6回実施した。提示したスライドを付録Fに示す。

### 5.7.3 収集するデータ

口頭発話を録音し、その回答内容を文字起こしした。

### 5.7.4 発話解釈の推論形式の導出

解釈選択課題で選択された解釈の推論形式は、多数回答課題で導出した推論形式をそのまま引用した。これは、解釈選択課題では、多数回答課題で生成された候補の中から選択するため、新たな推論形式のラベル付けは不要であるためである。

### 5.7.5 発話解釈の推論形式の導出対象

解釈選択課題では、6刺激すべてに対して1つずつ解釈が選択されるため、参加者1人あたり6個の推論形式が分析対象となった。なお、本課題は多数回答課題で回答した内容から選択しているため、無回答はなかった。

## 5.8 個人特性

### 5.8.1 目的

個人差測定のための目的は、発話解釈に影響しうる個人差を測定し、推論形式の使用頻度や解釈選択との関連を検討することである。

### 5.8.2 測定した心理測定尺度

本研究では、以下の心理測定尺度を測定した。

- (1) 短縮版日本語版 Big Five (並川 et al., 2012)
- (2) 日本語版自閉症スペクトラム指数(AQ-J)(コミュニケーションの下位1尺度) (若林 et al., 2004)
- (3) コミュニケーション・スキル尺度 ENDCOREs (解読力と他者受容の下位2

尺度) (藤本・大坊, 2007)

(4) 認知的欲求尺度 (神山・藤原, 1991)

(5) 多次元共感性尺度 (鈴木・木野, 2008)

(6) 認知的完結欲求尺度 (決断性と予測可能性に対する選好の下位 2 尺度) (鈴木・桜井, 2003)

心理測定尺度の測定は、次のリッカート尺度で測定した。

短縮版 Big Five, 認知的欲求尺度, 多次元共感尺度: 5 件法

- あてはまる
- ややあてはまる
- どちらともいえない
- ややあてはまらない
- あてはまらない

日本語版自閉症スペクトラム指数(AQ-J): 4 件法

- あてはまる
- どちらかといえばあてはまる
- どちらかといえばあてはまらない
- あてはまらないの,

コミュニケーション・スキル尺度 ENDCOREs: 7 件法

- かなり得意
- 得意
- やや得意
- ふつう
- やや苦手
- 苦手
- かなり苦手

認知的完結欲求尺度: 6 件法

- 非常にあてはまる
- あてはまる
- ややあてはまる
- ややあてはまらない

- あてはまらない
- 全くあてはまらない

心理測定尺度の測定には Google Form(以降, Form)を用いた. 被験者は Form が表紙された画面を見て, マウスを使用して心理測定尺度の項目に回答した. 心理測定尺度の指示項目は縦に並んでいる. Form での回答はプルダウンの選択肢で年齢, 性別, 心理測定尺度, ダミー質問 2 問を回答させた. なお, 心理測定尺度の回答には尺度に合わせたプルダウンの選択肢を設けた. ダミー質問は, 心理測定尺度回答中の注意力チェックとして, 「この質問にはあてはまるを選択してください」と教示した.

## 5.9 半構造化インタビュー

### 5.9.1 目的

半構造化インタビューの目的は, 発話解釈の生成および選択に関する参加者の主観的な方略を質的に把握することである. 特に, 量的データ(推論形式の頻度など)では捉えきれない, 参加者の意識的な判断基準や, 実験課題に対する感想を収集する. インタビューは, 実験終了後に口頭で実施し, 参加者の回答を録音した.

### 5.9.2 質問内容

インタビューの質問内容は以下の 2 点である.

1. 発話の解釈を考える際, 何に注目をしながら意図を考えましたか?
2. 設問 3 (解釈選択課題) では, 適切なものを 1 つ選択する課題でしたが, 何を基準に適切だと判断したのかを教えてください.

一つ目は, 参加者が, 状況, 関係性, 一般知識, 相手の感情など, どの情報を重視したかを把握する. 二つ目は, 参加者が, 「適切さ」をどのように定義しているか(状況整合性, 対人配慮, 一般的期待など)を把握する.

## 5.10 統計手法

本研究では, 以下の統計分析を実施した. なお, 分析には R(version 4.5.2)を使用した.

### 5.10.1 実験刺激の操作検証

本研究では, 実験刺激が想定どおり理解容易度の順序で操作できていたかを

検証するため、直感回答課題における反応時間を操作チェックの代理指標として用いて6刺激(A~F)で実施する。ただし、本課題では制限時間を15,000[ms]と設定しており、制限時間内に回答が得られなかった試行をタイムアウトとして扱う。タイムアウトは反応時間が15,000[ms]以上で、上限のため観測できなかった打ち切りである。

この点を考慮して、実験刺激の操作チェックは、反応時間とタイムアウトが発生した試行の割合(以降、タイムアウト率)を判別指標として用いる。理由は二つあり、一つ目は刺激の理解容易度が難しければ、参加者は制限時間内に解釈へ到達できずタイムアウトが増加するからだ。二点目は、タイムアウトが多い刺激ほど「遅すぎて時間内に回答できなかった試行」が多く除かれ、偏りがうまれるからだ。例として理解容易度が低い刺激(E・F)ほど「比較的早く回答できた試行のみ」に偏ってしまい、実験刺激操作の検証として根拠が弱くなる。以上を踏まえ、実験刺激の操作チェックは、6刺激(A~F)ごとのタイムアウト率を算出し、刺激間の差を示す。

#### 5.10.2 仮説の検証

推論形式が刺激によってどのように変化するかを、使用回数と使用比率の視点から検討する。そのため、各課題においてどのように推論形式が用いられたかを集計し、刺激別での分布と傾向を見る。

推論形式の使用頻度は、多数回答課題において各刺激で生成された解釈の説明に対して同定した各推論形式の数と割合で求める。

#### 5.10.3 心理測定尺度

各心理測定尺度について、記述統計量(平均値、標準偏差、最小値、最大値)を算出した。また、内的整合性の評価には、Cronbachの $\alpha$ 係数およびMcDonaldの $\omega$ 係数を算出した。

#### 5.10.4 個人特性と推論形式の関係

本研究では、発話解釈で使用する推論形式がどのような個人特性を含んでいるか傾向を分析する。なお本分析では全体の相関を確認はするが、詳細な関係については統計回数の観点から絞って分析を実施し、他は探索的分析として解釈する。詳細は6.5節に示す。

## 第6章 結果

5章では本研究の仮説を検証する実験の詳細を述べた。実験では直感回答課題で発話解釈にかかる反応時間と推論形式。多数回答課題では解釈の生成数と推論形式。解釈選択課題では解釈選択の推論形式を観察した。

本章では、本実験の結果について述べる。分析には反復測定に対応したモデルを採用する。6.1節では、実験データの正規性を確認し、実験刺激がAからFに向けて理解容易度が低くなるように操作できていたかの妥当性を検証する。6.2節では Think Aloud 法で取得した回答からどのように推論形式を判別したのかを参加者から得た回答を基に示す。6.3節では多数回答課題の回答を用いて推論形式の使用回数・比率が刺激別で差があるかを検証する。6.4節では、個人特性と推論形式の関係を検証する。6.5節では、各課題の推論形式の頻度と個人特性の関係を述べる。

### 6.1 実験刺激の操作チェック

#### 6.1.1 各実験刺激の反応時間分布

実験刺激A～Fで計測した反応時間の記述統計量を算出したものを表6.1.1に示す。なお、反応時間が15,000[ms]を超えたものはタイムアウトとして処理した。図6.1.1に刺激別の反応時間分布を示す。平均反応時間（折れ線）は、AとBで短く、C以降で増加しEで高い傾向を示した。したがって、刺激間で処理負荷が異なる可能性が示唆される。また、各刺激において箱ひげ図の上部が長く、上方向にプロットが多く散らばっているため、反復時間は右裾の長い歪みのある分布であることが示唆される。本分析では正規性がないものと判断し、適切な統計検定を使用する。次節では、この差が統計的に有意であるかを検証する。

表 6.1.1 各刺激の反応時間の記述統計量

刺激	n	M	SD	Md	Min	Max
A	38	6985	2650	6544	2637	15000
B	38	6270	2576	5746	2420	15000
C	38	8344	2740	7886	3015	15000
D	38	9640	3715	9454	3316	15000
E	38	11241	3382	11120	6084	15000
F	38	10138	3374	10150	4566	15000

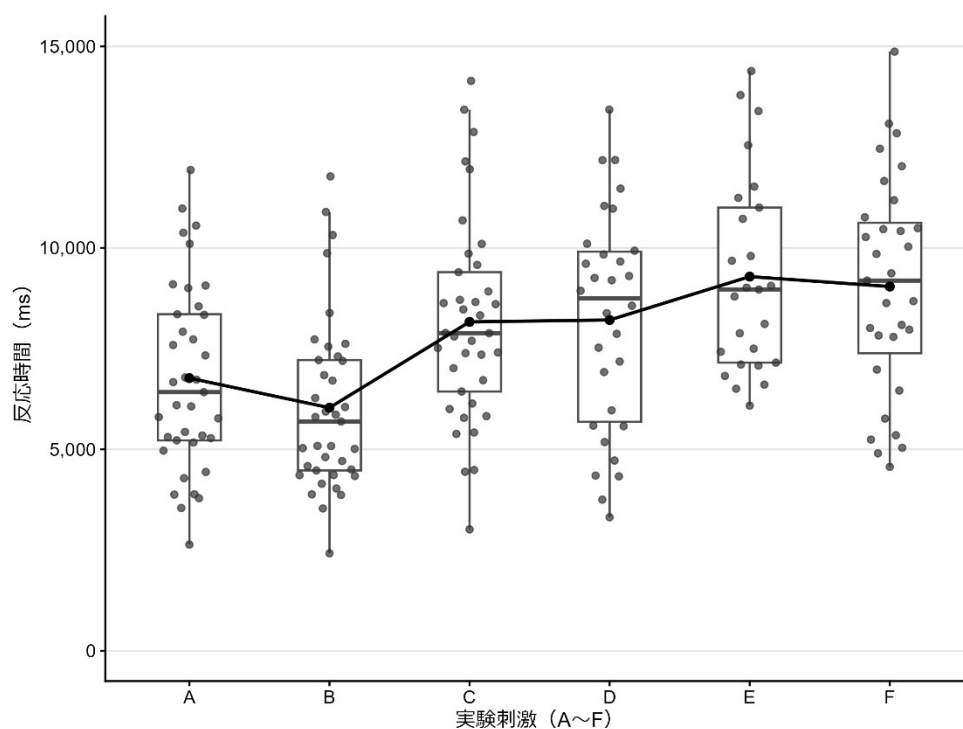


図 6.1.1 各刺激の反応時間分布

縦軸が反応時間，横軸が各刺激，プロットは参加者，折れ線は平均

### 6.1.2 各実験刺激の反応時間差の検証

各刺激の反応時間に差があるかを検証するため，フリードマン (Friedman) 検定検定を実施した。フリードマン検定とは，反復測定をしており，従属変数が順序尺度や正規性が怪しい場合でも使用でき，条件に差があるかを見る検定である。本分析には欠損がないため参加者数全員の回答を解析対象とした。その結果，反応時間は刺激間で差が中効果で認められた  $\chi^2(5)=62.54$ ,  $p<.001$  と 5%水準で有意であった (表 6.1.2)。

表 6.1.2 各刺激の反応時間差の検定

N	刺激	検定量	自由度	p 値	Kendall's W
38	6	62.54	5	<.001	0.329

フリードマン検定により各刺激で有意差を確認されたため，次に事前に定めた理解容易度が正しく機能しているかをウィルコックスン (Wilcoxon) の符号付順位検定を実施した。ウィルコックスンの符号付順位検定とは，反復測定をしており，従属変数が順序尺度や正規性が怪しい場合でも使用でき，条件に差がある

かを見る検定である。また、多重比較に伴う検定を行うと、偶然に有意になる確率が増えるために p 値は Holm 法で補正した。Holm 法とは n 回の検定を実施し、得られた p 値を小さい順に並べて、小さい順に多重比較の回数 - 1 を p 値に積算し補正を行う。例を挙げると、本分析では 3 回の検定を実施して小さい順に 3 倍、2 倍、1 倍と p 値を補正する。本ペアは図 6.2.1 を基に理解容易度の代表値の比較で検証した。多重比較の結果、B-D ( $p < .001, z = -4.62, r = .76$ ), D-E ( $p = .003, z = -2.12, r = .37$ ), B-E ( $p < .001, z = -5.11, r = .83$ ) と有意となり Holm 補正後も有意差が確認された (表 6.1.3)。ここで、 $r = |z|/\sqrt{N}$  は効果量である。以上の結果より、理解容易度は高中低の順に遅くなっていることが統計的に認められた。

**表 6.1.3 理解容易度の差の検証**

多重比較	検定量	p: 補正前	p: 補正後
B vs D	45	< .001	< .001
D vs E	162	0.03	0.03
B vs E	18	< .001	< .001

### 6.1.3 各刺激のタイムアウト率分布

実験刺激 A~F で計測した反応時間から算出したタイムアウト率の記述統計量を表 6.1.4 に示す。図 6.1.5 に刺激別のタイムアウト率分布を示す。記述統計から、6 つの刺激は理解容易度高群 (ABC) と低群 (DEF) の 2 群に分かれることが示された。特に刺激 E は最も高いタイムアウト率を示し、約 3 割の参加者が回答できなかった。次節では、この差が統計的に有意であるかを検証する。

**表 6.1.4 各刺激のタイムアウト率の記述統計量**

刺激	N	タイムアウト数	タイムアウト率
A	38	1	2.63
B	38	1	2.63
C	38	1	2.63
D	38	8	21.1
E	38	13	34.2
F	38	7	18.4

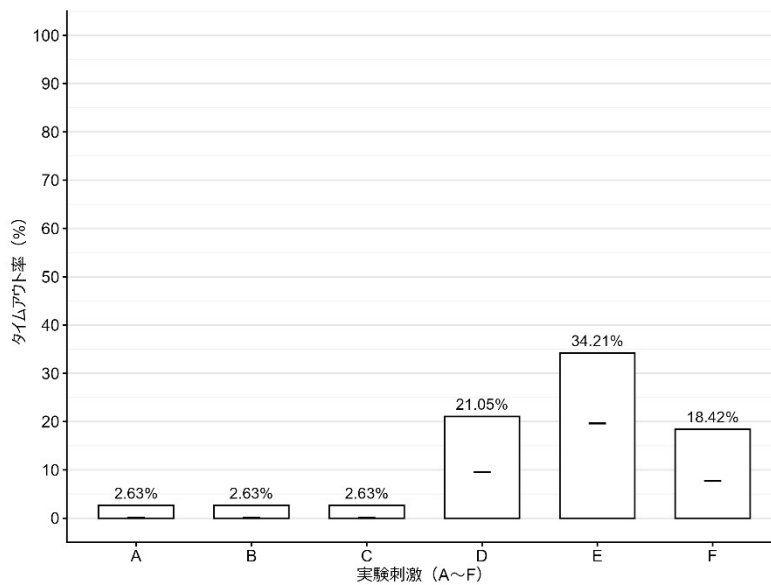


表 6.1.5 刺激別のタイムアウト率分布

#### 6.1.4 タイムアウト率から見た各刺激の差

刺激間のタイムアウト率に差があるかを検討するため、コ克兰の Q 検定を実施した。コ克兰の Q 検定とは、反復測定をしており、従属変数が二値であるとき、条件間に差があるかを見る検定である。本研究では刺激 6 つを参加者は経験しており、反応時間についてタイムアウト有無を指標とし、各刺激でタイムアウト率が異なるかを検証した。なお、データの欠損がなかったため  $n=38$  を解析対象とした。結果は、6 つの刺激間のタイムアウト率に  $Q(6) = 30.95$ ,  $p < .001$  と有意差が認められたが、Kendall's  $W = .163$  と効果量は小さく、刺激間の差よりも個人差の方が大きいことが示唆された(表 6.1.6)。以上より対立仮説である刺激別に理解容易度の差があることが支持された。

表 6.1.6 各刺激に関するタイムアウト率の比較

N	条件数(k)	検定量	自由度	p 値	Kendall's W
38	6	30.95	5	< .001	0.163

コ克兰の Q 検定によりタイムアウト率に刺激間に有意差が確認されたため、次に事前に定めた理解容易度が正しく機能しているかマクネマー (McNemar) 検定を実施した。マクネマー検定とは、反復測定をしており、従属変数が二値であるときに用いる検定である。また、単なる割合ではなく「同じ人が A では 0, B で 1 に変わった」のか変化に着目して差を見る。また、多重比較に伴う検定を

行くと、偶然に有意になる確率が増えるために p 値は Holm 法で補正した。本ペアは図 6.2.5 を基に理解容易度の代表値の比較で検証した。結果、理解容易度：高(B)と中(D)の間、高(B)と低(E)と中(C)と中(D)の間に有意差が確認され、中(D)と低(E)の間では有意な差が無かった。これは C が理解容易度に入っており、想定していた実験操作が行われていなかった。結果を表 6.1.7 にまとめた。

表 6.1.7 理解容易度の差の検証

比較ペア	N	B 値	C 値	検定量	自由度	p: 補正前	p: 補正後
B vs D	38	3	21	5.14	1	0.016	0.047*
D vs E	38	21	34	1.45	1	0.227	0.227
B vs E	38	3	34	8.64	1	0.002	0.007*
C vs D	38	3	21	5.14	1	0.016	0.047*

c > b の場合、対象刺激の方がタイムアウトは多く、より困難であることを意味する

#### 6.1.5 操作チェックのまとめ

反応時間およびタイムアウト率の 2 指標から、実験刺激の理解容易度を検証した結果、以下が明らかとなった。反応時間では、刺激間に有意差が認められ ( $\chi^2(5) = 62.54, p < .001$ )。事後検定では、すべてのペアで有意差が確認された。

タイムアウト率では、ABC と DEF の 2 群が観察され、刺激間に有意差が認められた ( $Q = 30.69, p < .001$ )。事後検定では B vs D ( $p = .047$ )、B vs E (.007)、C vs D (.047) 間で有意差が認められ、D vs E (.227) では有意差が認められなかった。

以上の結果より、理解容易度の操作は部分的に成功した。高群(A,B)および低群(E,F)では明確に区別されたが、中群(C,D)の位置付けは不明確であった。反応時間では刺激 D と E に有意差が認められた ( $p = .003$ ) が、タイムアウト率では有意差が見られず、両指標で異なる傾向を示した。以降の分析では、理解容易度の 3 水準は完全には分離しない可能性があるため、以降は刺激 A-F の比較を主分析として提示する。

## 6.2 Think Aloud 法で取得した回答の推論形式判別

### 6.2.1 推論形式の判別例

多数回答課題で生成された解釈は、推論形式分類基準(表 4.4.1)を用いて推定(アブダクション)、規則(演繹)、認知的共感の3つの推論形式に分類した。なお、直感回答課題および解釈選択課題の判別に関するの詳細は 5.5.6 節および 5.7.4 節に示した。以下に3つの推論形式をどのように判別したかを参加者の回答を用いて具体例を示す。下線は回答者の考えた解釈、斜体は解釈の説明を示す。

#### [推定]

・刺激：B「レストランでメニューを訪ねている」

・回答：「カロリー高いですねとはいつつも、ちょっと食べてみたいから別に頼んでみてもいいかもと考えた」

・判別ポイント：

➤ 「自分も頼んでもいいかも」と可能性を示しているため推定と判定

・刺激：F「友人の誕生日プレゼントを探している」

・回答：「レイカは花束についての花言葉の意味を知っているから、プレゼントとしてそのまま使えろと考えている」

・判別ポイント：

➤ 自分の伝えたい思いが花束に含まれてると仮定しているため推定と判定

#### [規則]

・刺激：B「レストランでメニューを訪ねている」

・回答：「この揚げ物のカロリーが高いことを気にしているので、カロリーの高いものはあまり好まない」

・判別ポイント：

➤ カロリーを気にしている人はカロリーの高いものは好まないという一般規則から規則と判定

・刺激：F「友人の誕生日プレゼントを探している」

・回答：「もしかしたら渡す友達は、花束が嫌いっていうのをレイカだけ知っていて、やめといた方がいいんじゃない?」

・判別ポイント：

- 嫌いなものを渡すのは良くないという一般規則から規則と判定

#### [認知的共感]

- ・刺激：B「レストランでメニューを訪ねている」

- ・回答：「シンプルに揚げ物の中でも、このレストランのこの揚げ物のカロリーがとても高かったので、そこにびっくりしてあなたにこの情報を共有したいという意図でこの発言をしている。」

- ・判別ポイント：

- 「この情報を共有したい」と視点取得を示しているため認知的共感と判定

- ・刺激：F「友人の誕生日プレゼントを探している」

- ・回答：「そのものが使われているねとは、花束の特徴をぱっとこう抜きだしたことに対して、この花束は結構特徴的で、すごくこの花束がいいなあと、ポジティブにとらえている。」

- ・判別ポイント：

- 「すごくこの花束がいいな」と心的状態を示した視点取得を行っているため認知的共感と判定

#### 6.2.2 推論形式使用回数・比率の傾向

多数回答課題（Q2）における各刺激の推論形式使用回数および比率の記述統計量を表 6.2.1 に示す。また、各刺激の推論形式使用の分布を図 6.2.1、推論形式比率の分布を図 6.2.2 にそれぞれ示す。

推論形式使用回数の特徴として、理解容易度の高い刺激から刺激 C まで減少し、刺激 F に向かって増加傾向を示した。また、どの刺激でも推定の使用回数が多く、次に規則・認知的共感の順で減少傾向にある。

推論形式使用比率の特徴として、理解容易度の低い刺激 F で推定が特に高い。また、中程度の刺激 D で規則が高い傾向が見られた。一方、認知的共感はすべての刺激で低く、特に刺激 E で最も少なかった。

表 6.2.1 各刺激の推論形式使用回数・比率全体の傾向

刺激	総使用回数	推定		規則		認知的共感	
		使用回数	比率	使用回数	比率	使用回数	比率
A	166	104	62.7	53	31.9	9	5.4
B	174	107	61.5	59	33.9	8	4.6
C	119	82	68.9	20	16.8	17	14.3
D	147	77	52.4	66	44.9	4	2.7
E	125	74	59.2	49	39.2	2	1.6
F	126	99	78.6	13	10.3	14	11.1

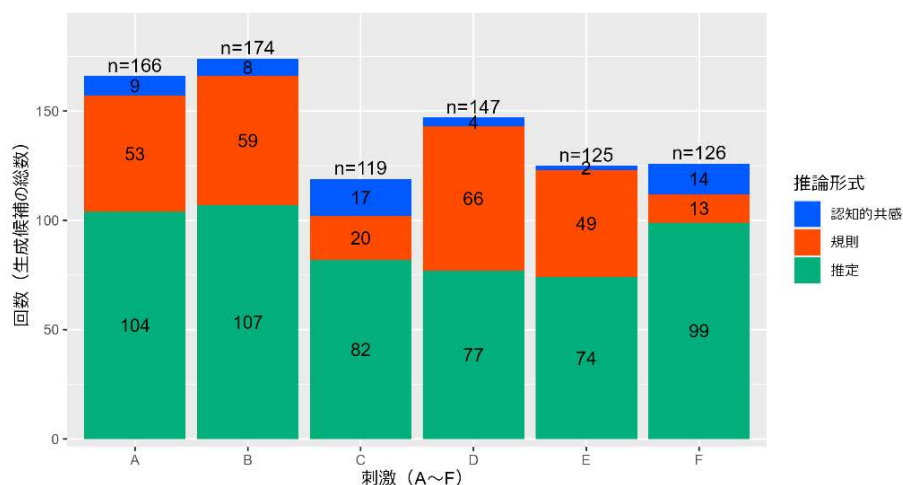


図 6.2.1 各刺激の推論形式使用回数分布

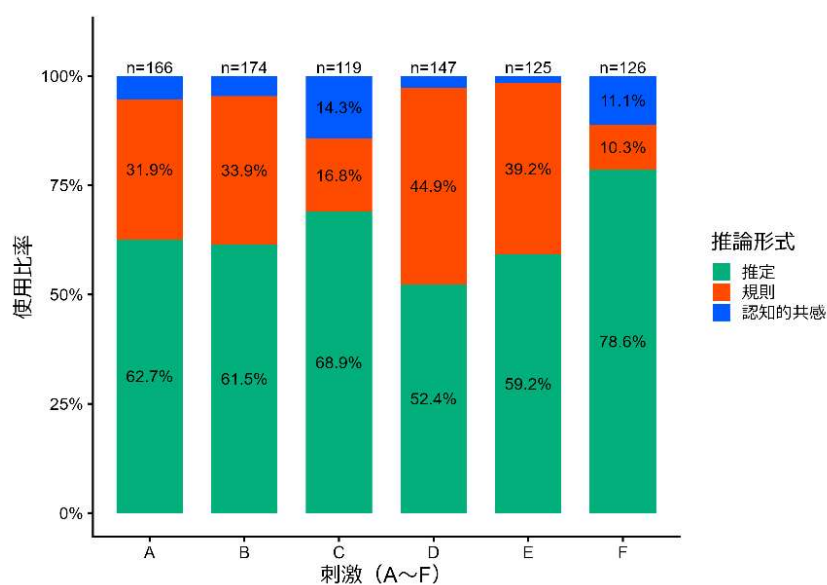


図 6.2.2 各刺激の推論形式使用比率分布

## 6.3 理解容易度が推論形式使用回数に与える影響

### 6.3.1 記述統計量

多数回答課題における各刺激の推論形式使用回数の記述統計量を表 6.3.1 に示す。また、推論形式ごとの使用回数分布を図 6.3.1, 6.3.2, 6.3.3 に示す。

表 6.3.1 各刺激の推論形式使用回数の記述統計量

刺激	推論形式	M	SD	MD	MAX	MIN
A	推定	2.74	1.52	2.5	7	0
A	規則	1.39	1.24	1	6	0
A	認知的共感	0.24	0.49	0	2	0
B	推定	2.82	1.29	2.5	6	1
B	規則	1.55	1.20	1	5	0
B	認知的共感	0.21	0.41	0	1	0
C	推定	2.16	1.37	2	6	0
C	規則	0.53	0.83	0	4	0
C	認知的共感	0.45	0.60	0	2	0
D	推定	2.03	1.10	2	4	0
D	規則	1.74	1.48	2	6	0
D	認知的共感	0.11	0.31	0	1	0
E	推定	1.95	1.39	2	6	0
E	規則	1.29	1.14	1	4	0
E	認知的共感	0.05	0.23	0	1	0
F	推定	2.61	1.20	2.5	5	0
F	規則	0.34	0.71	0	3	0
F	認知的共感	0.37	0.59	0	2	0

記述統計量より次の傾向を確認した。推定は、どの課題でも最も使用されている ( $M = 2.74$ ,  $SD = 1.52$ )。規則は、推定よりは使用が少ない ( $M = 2.74$ ,  $SD = 1.52$ )。認知的共感は、3つの中で最も使用回数が少ない ( $M = 2.74$ ,  $SD = 1.52$ )。

次に推論形式ごとの分布を確認した。推定・規則・認知的共感は、すべて刺激間で個人差が大きい。特に規則と認知的共感では、参加者が 0 回または少数回の使用に集中しており、正規分布から大きく逸脱していた。本分析では正規性が無いものと判断し、適切な統計検定を使用する。次節では、この差が統計的に有

意であるかを検証する。

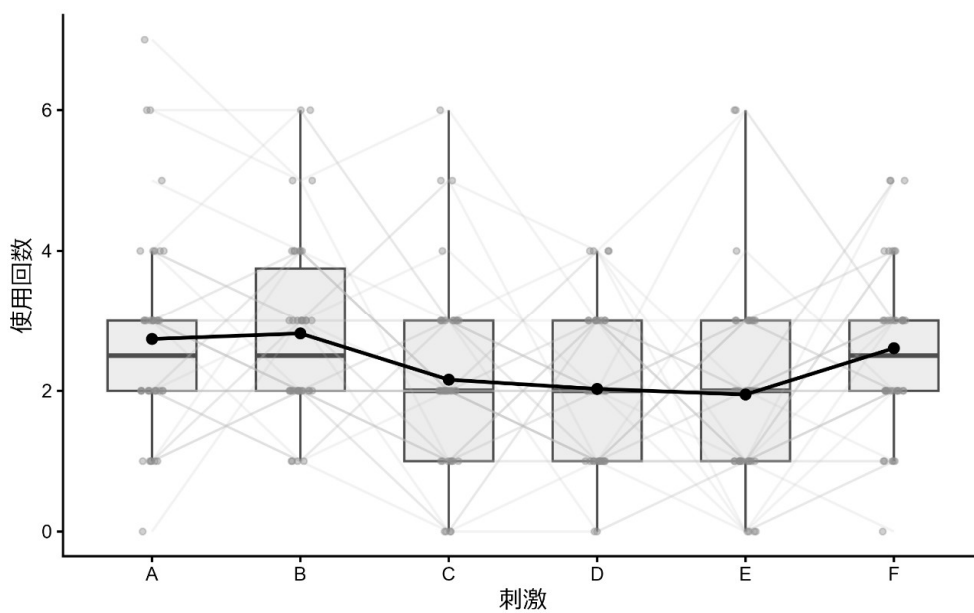


図 6.3.1 多数回答課題 (Q2) の「推定」使用回数分布  
 プロットは参加者, 折れ線は平均値

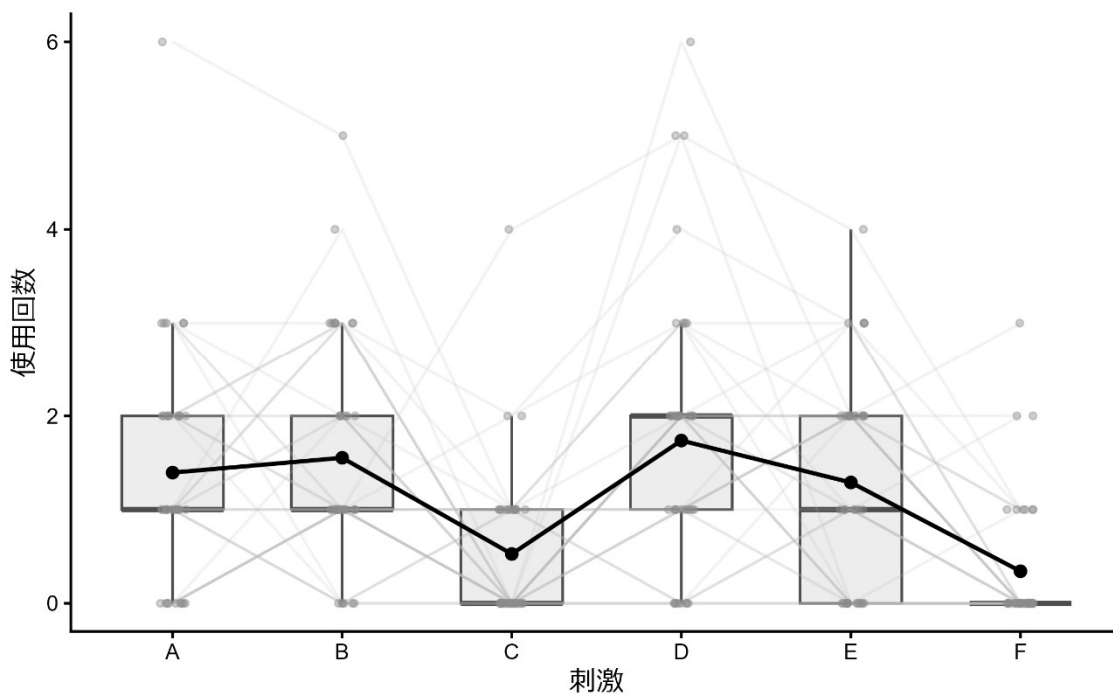


図 6.3.2 多数回答課題の「規則」使用回数分布  
 プロットは参加者, 折れ線は平均値

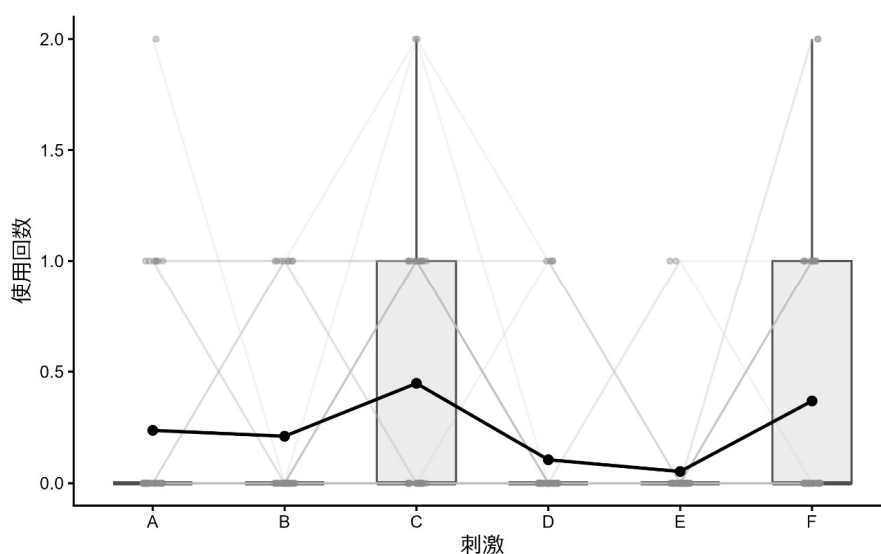


図 6.3.3 多数回答課題の「認知的共感」使用回数分布  
 プロットは参加者，折れ線は平均値

### 6.3.2 各刺激の推論形式使用回数の差

刺激間で推論形式の使用回数に差があるかを検証するため，推論形式ごとにフリードマン(Friedman)検定を実施した．フリードマン検定とは，反復測定をしており，従属変数が順序尺度や正規性が怪しい場合でも使用でき，条件に差があるかを見る検定である．結果，推論形式ごと（推定： $\chi^2(5) = 18.57, p < .001$ ，規則： $\chi^2(5) = 54.31, p < .001$ ，認知的共感： $\chi^2(5) = 18.89, p < .001$ ）と 5%水準で刺激間に統計的に有意な差が認められた（表 6.3.2）．

表 6.3.2 各推論形式使用回数の検定結果集計

推論形式	検定量	自由度	p 値	Kendall's W
推定	18.57	5	< .001	0.098
規則	54.31	5	< .001	0.286
認知的共感	18.89	5	< .001	0.099

### 6.3.3 理解容易度の推論形式使用回数の差

フリードマン検定により刺激間に有意差が確認されたため，次にどの刺激のペア間に差があるのかをウィルコックソンの符号付順位検定で検証した．この検定では理解容易度で推論形式ごとに差があるかの検証も調査する．ペアは図 6.3.1～3 を参考に仮説の検定および目視で差がありそうな部分を検証した．なお，認知的共感の使用回数が少なかったため分析から除外した．

**表 6.3.3 各刺激の推論使用回数の有意差**

比較	検定量	p:補正前	p:補正後
A vs F	238	0.65	0.68
B vs C	327	0.02	0.06
C vs E	197	0.34	0.68
B vs E	341	< .001	0.03
F vs E	322	0.02	0.06

**表 6.3.4 各刺激の規則使用回数の有意差**

比較	検定量	p:補正前	p:補正後
A vs F	397	< .001	< .001
B vs C	379	< .001	< .001
C vs D	18	< .001	< .001
C vs E	22.5	< .001	< .001
B vs E	255	.407	.407
F vs E	20	< .001	< .001

推定は B vs C ( $p = .002$ , 補正後  $p = .005$ ), B vs E ( $p = < .001$ , 補正後  $p = .003$ ), F vs E ( $p = .002$ , 補正後  $p = .005$ ) で補正後にも有意差が見られた。一方, 理解容易度の端点である A vs F ( $p = .65$ , 補正後  $p = .65$ ) では有意差が見られなかった (表 6.3.3)。

規則は全てのペア (AF : ( $p = < .001$ , 補正後  $p = < .001$ ), BC : ( $p = < .001$ , 補正後  $p = < .001$ ), CD : ( $p = < .001$ , 補正後  $p = < .001$ ), FE : ( $p = < .001$ , 補正後  $p = < .001$ )) で補正後も有意差が見られた。

以上より, 推定の使用回数は一部のペアで有意差が見られたが, 理解容易度の順序に沿った傾向は確認されなかった。一方, 規則の使用回数はすべてのペアで有意差が見られ, 刺激間で有意な差があることが示された (表 6.3.4)。

## 6.4 理解容易度が推論形式使用比率に与える影響

### 6.4.1 記述統計量

多数回答課題における各刺激の推論形式使用比率の記述統計量を表 6.4.1 に示す。また, 推論形式ごとの使用比率分布を図 6.3.1, 6.3.2, 6.3.3 に示す。

記述統計量より次の傾向を確認した。推定は, 刺激 F で特に使用比率 ( $M =$

0.78) が高く、他の刺激より推定が使用しやすい刺激だと考えられる。規則は、刺激 D で最も使用比率が高く (M = 0.43)、刺激 F が極端に低い (M = 0.10) ことが分かる。これは推定による影響であると考えられる。認知的共感は、すべての刺激で 20%以下であった。また、3つの推論形式はいずれも理解容易度の順序とは一貫していないことが示された。

次は分布を確認する。推定では、個人差が大きく、刺激 C・E では箱が上部に偏る分布を示した。規則は、刺激 C・F で箱が 0 付近に強く圧縮され、多くの参加者が規則をほとんど使用していない分布を示した。認知的共感はすべての刺激で箱が 0 付近であり、一部の参加者のみが使用する分布を示した。

表 6.4.1 各刺激の推論形式使用比率の記述統計量

刺激	推論形式	M	SD	MD	MAX	MIN
A	推定	0.61	0.24	0.63	1.00	0
A	規則	0.31	0.24	0.27	0.75	0
A	認知的共感	0.08	0.19	0	1.00	0
B	推定	0.62	0.22	0.67	1.00	0.25
B	規則	0.33	0.21	0.33	0.75	0
B	認知的共感	0.05	0.10	0	0.33	0
C	推定	0.65	0.31	0.67	1.00	0
C	規則	0.17	0.25	0	1.00	0
C	認知的共感	0.18	0.27	0	1.00	0
D	推定	0.54	0.29	0.50	1.00	0
D	規則	0.43	0.29	0.50	1.00	0
D	認知的共感	0.03	0.09	0	0.33	0
E	推定	0.60	0.34	0.50	1.00	0
E	規則	0.39	0.34	0.45	1.00	0
E	認知的共感	0.01	0.06	0	0.33	0
F	推定	0.78	0.26	0.92	1.00	0
F	規則	0.10	0.20	0	0.67	0
F	認知的共感	0.12	0.22	0	1.00	0

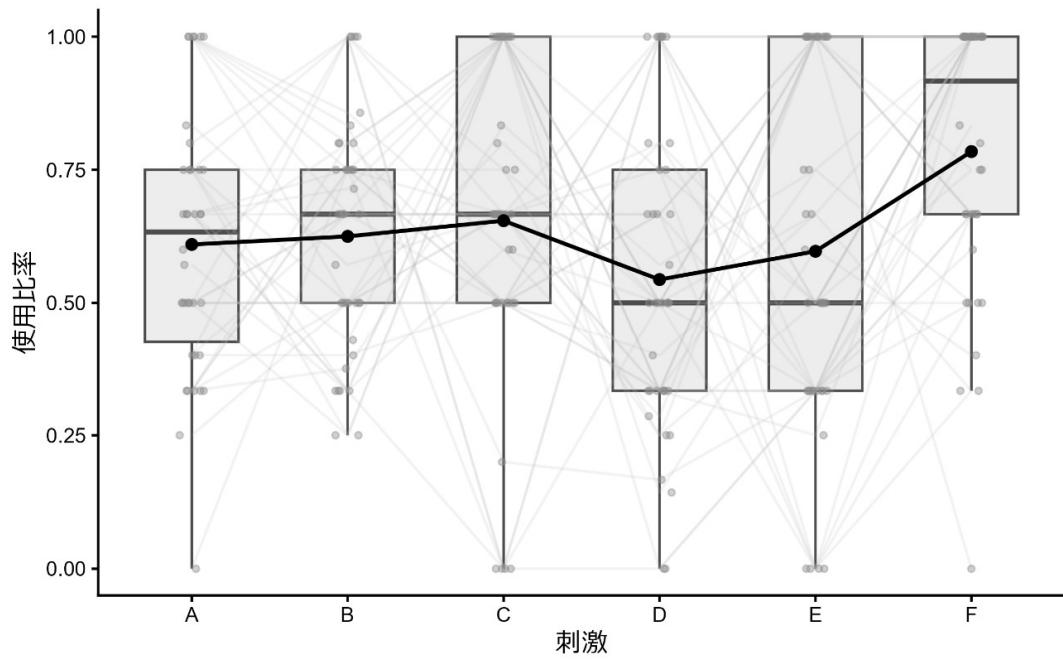


図 6.4.1 多数回答課題の「推定」使用比率分布  
 プロットは参加者, 折れ線は平均値

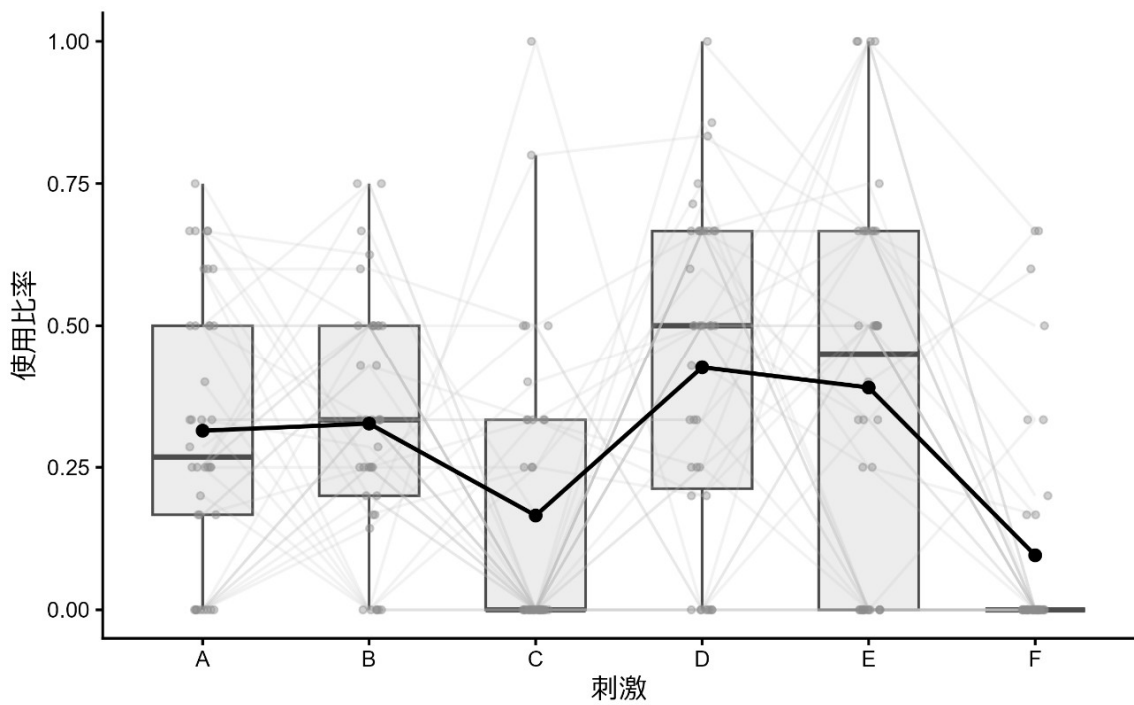


図 6.4.2 多数回答課題の「規則」使用比率分布  
 プロットは参加者, 折れ線は平均値

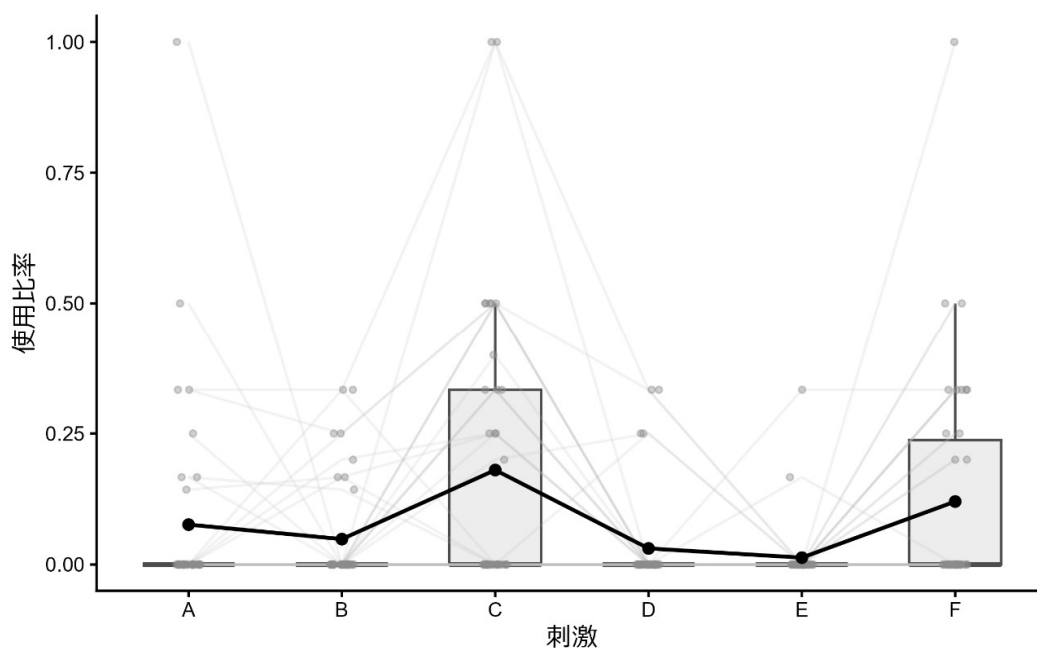


図 6.4.3 多数回答課題の「認知的共感」使用比率分布  
 プロットは参加者，折れ線は平均値

#### 6.4.1 各刺激の推論形式使用比率の差の検証

刺激間で推論形式の使用比率に差があるかを検証するため，推論形式ごとにフリードマン(Friedman)検定を実施した．本分析には欠損がなかったため参加者数全員の回答を解析対象とした．表 6.4.2 に推論形式別のフリードマン検定の結果を示す．結果，全ての推論形式において，刺激間に統計的に有意な差が認められたが，Kendall's W = .098 (推定)，.214 (規則)，.115 (認知的共感) と効果量は小さく，刺激間の差よりも個人差の方が大きいことが示唆された．以上より対立仮説である刺激間で推論形式の使用頻度に差があることが支持された．

表 6.4.2 各推論形式使用比率のフリードマン検定の集計

推論形式	検定量	自由度	p 値	Kendall's W
推定	18.63	5	0.002	0.098
規則	40.64	5	< .001	0.214
認知的共感	21.93	5	< .001	0.115

#### 6.4.2 理解容易度の推論形式使用比率の差の検証

フリードマン検定により刺激間に有意差が確認されたため、次にどの刺激のペア間に差があるのかをウィルコクソンの符号付順位検定で検証した。この検定では理解容易度で推論形式ごとに差があるかの検証も調査する。ペアは図 6.4.1～3 を参考に仮説の検定および目視で差がありそうな部分を検証した。なお、認知的共感は使用回数が少なくため分析から除外した。

**表 6.4.3 各刺激の「推定」使用の差**

比較	検定量	p:補正前	p:補正後
B vs D	321	0.15	0.61
D vs E	223	0.44	1.00
B vs E	254	0.91	1.00
C vs D	377	0.09	0.43
C vs E	286	0.46	1.00
F vs A	500	< .001	0.02
F vs B	419	< .001	0.03
F vs C	295	0.04	0.22
F vs D	518	< .001	< .001
F vs E	346	< .001	0.04

**表 6.4.4 各刺激の「規則」使用確率の差**

比較	検定量	p:補正前	p:補正後
B vs D	205	0.11	0.44
D vs E	259	0.59	1.00
B vs E	236	0.42	1.00
C vs D	74	< .001	< .001
C vs E	54	< .001	< .001
F vs A	64	< .001	< .001
F vs B	61	< .001	< .001
F vs C	72	0.35	1.00
F vs D	29	< .001	< .001
F vs E	17	< .001	< .001

推定の使用比率は、F vs A ( $p < .001$ , 補正後  $p = .002$ ), F vs B ( $p < .001$ , 補正後  $p = .003$ ), F vs D ( $p < .001$ , 補正後  $p < .001$ ), F vs E ( $p < .001$ , 補正後  $p = .004$ ) で有意差が見られた。

推定の使用比率は、F vs A ( $p < .001$ , 補正後  $p < .001$ ), F vs B ( $p < .001$ , 補正後  $p < .001$ ), F vs D ( $p < .001$ , 補正後  $p < .001$ ), F vs E ( $p < .001$ , 補正後  $p < .001$ ) で有意差が見られた。

以上より、刺激 F は他の刺激と比較して、推定の使用比率が有意に高く、規則の使用比率が有意に低いことが示された。特に F vs D で最も大きな差が見られた。

## 6.5 個人特性が推定形式に与える影響

### 6.5.1 記述統計量

本調査で使用した各心理測定尺度の記述統計量を求めた (表 6.5.1)。Big Five の 5 因子は、すべての因子で一定の個人差が見られた ( $SD = 3.97-6.22$ )。自閉症スペクトラムのコミュニケーション因子でも、同様に ( $SD = 2.59$ ) 個人差が大きいことが示された。なお、自閉症スペクトラムは 1 つの因子のみでは解釈が難しいため、分析から除く。認知的欲求では、得点範囲が 47-97 点と高低に大きな個人差が見られた。多次元共感性尺度では、すべての因子で一定の個人差が見られた ( $SD = 3.12-4.49$ )。また、自己指向的反応は ( $SD = 3.12$ ) と他の因子より低かった。認知的完結欲求は得点範囲が 5-24 点と個人差が大きかった。

以上の結果より、本実験の参加者は各心理測定尺度において一定の個人差を示し、推論形式の使用との関係を検討するための十分な変動が確認された。

次に各心理測定尺度で得られたデータに内的整合性を評価するためにクロムバック (Cronbach) の  $\alpha$  係数とマクドナルド (McDonald) の  $\omega$  係数を求めた (表 6.5.2)。クロムバックの  $\alpha$  係数とは、尺度を構成する項目間の内的整合性を評価する指標であり、.70 以上が望ましいとされる。マクドナルドの  $\omega$  係数とは、すべての項目が同じ重みを持つと仮定する  $\alpha$  係数と異なり、項目ごとの重みの違いを考慮した信頼性指標である。

信頼性分析の結果、すべての尺度でクロムバックの  $\alpha$  係数は .570 から .878 の範囲、マクドナルドの  $\omega$  係数は .625 から .890 の範囲にあった。多次元共感性尺度の自己指向的反応因子 ( $\alpha = .570$ ,  $\omega = .625$ ) がやや低い値を示したものの、その他の尺度・因子はすべて  $\alpha > .70$  の基準を満たし、信頼性が確認された。

表 6.5.1 心理測定尺度の記述統計量 (N=38)

心理測定尺度	M	SD	MD	IQR	Min	Max
<b>短縮版日本語版 Big Five</b>						
外向性	18.0	4.71	18	7.75	7	26
誠実性	28.0	6.22	28	8.75	19	43
情緒不安定性	16.8	4.86	17	8	5	25
開放性	19.9	3.97	21	5.5	10	28
調和性	27.0	5.37	28	8.5	16	36
<b>自閉症 スペクトラム尺度</b>						
コミュニケーション	5.68	2.59	6	4	0	9
<b>コミュニケーション・ スキル尺度 ENDCOREs</b>						
解読力	18.5	4.47	19	5.75	8	28
他者受容	20.7	4.19	21	6	11	28
<b>認知的欲求</b>						
認知的欲求	75.2	15.0	80	19.3	47	97
<b>多次元共感性尺度 MES</b>						
被影響性	14.4	3.19	15	4.75	8	20
他者指向的反応	19.7	3.12	20	6	13	25
想像性	17.5	4.49	18	5.5	5	24
視点取得	19.7	4.06	20	4	11	25
自己指向的反応	13.0	3.12	13	3.75	4	18
<b>認知的完結欲求</b>						
決断性	23.8	7.09	23	6.5	11	43
予測可能性に対する 選好	13.8	5.30	14	8.5	5	24

表 6.5.2 心理測定尺度の信頼性分析 (N=38)

心理測定尺度	$\alpha$	$\omega$
短縮版日本語版 Big Five		
外向性	0.809	0.813
誠実性	0.821	0.827
情緒不安定性	0.767	0.785
開放性	0.692	0.697
調和性	0.815	0.825
自閉症スペクトラム尺度		
コミュニケーション	0.732	0.743
コミュニケーション・スキル尺度 ENDCOREs		
解読力	0.789	0.798
他者受容	0.727	0.761
認知的欲求		
認知的欲求	0.878	0.89
多次元共感性尺度:MES		
被影響性	0.862	0.869
他者指向的反応	0.699	0.733
想像性	0.765	0.791
視点取得	0.849	0.861
自己指向的反応	0.57	0.625
認知的完結欲求		
決断性	0.805	0.816
予測可能性に対する選好	0.847	0.859

### 6.5.2 個人特性が与える推論形式使用回数への影響

個人特性が、推論形式の使用回数にどのような傾向があるかを検証するため、各課題の推論形式使用回数でスピアマンの順位相関を実施した。スピアマンの順位相関とは、2つの変数を順位に変換し、順位間でどのような関係があるのかをみる無相関検定である。この指標では順番に注目しており、心理測定尺度で測った値が高い人ほどどのような推論方略で発話解釈を行っているのかを探索的に検証する。多重比較には Benjamini-Hochberg 法を用いた FDR 補正を行った。FDR 補正とは、多重比較による偶然小さい p 値が生じる問題に対して「有意と判定された結果（棄却した帰無仮説）」に含まれる誤って有意と判定した割合を特定の水準  $q$  以下（本実験では  $q < 0.05$ ）に抑える補正である。Benjamini-Hochberg 法は  $m$  個の  $p$  値を小さい順に並べ、並べた順番に応じて順位に合わせて合格ラインを作成する。合格ラインは（順位 ÷ 全検定数）\*  $q$  という式で計算する。そして、各  $p$  値がその順位の合格ラインより小さいか順に確認し、合格できたところまでを有意と判断する。

本分析は多数回答課題を対象とする。理由は、直感回答課題（Q1）および解釈選択課題（Q3）は回答数が1つであり、相関が正確に出ない可能性がある。また、認知的共感も使用回数が17回（比率：15%以下）であり、数が少ないため本分析では対象外とする。

心理尺度の選定は、2つの条件を満たすものを対象とした。一つ目は言語や非言語のコミュニケーションを適切に扱う技能を測定もの。二つ目は、他者の視点や心的状態を理解しようとする能力を測定するもの、を選択した。以上の条件を満たした本分析の対象を以下に示す。

- 推論形式：推定（アブダクション）、規則（帰納・演繹）
- 心理尺度：コミュニケーション・スキル尺度（ENDCOREs）、多次元共感性尺度、認知的欲求、認知的完結欲求

推定形式ごとのスピアマンの順位相関の結果を表 6.5.3 に示す。また、各推論形式および総回答数と有意な相関が出た個人特性（解読力および他者受容：ENDCOREs）の相関図を図 6.5.1, 6.5.2 に示す。

ENDCOREs の解読力および他者受容と規則の使用回数との間に有意な正の相関が見られた（ $\rho = 0.413$ , 補正後  $p = .010$ , 他者受容： $\rho = 0.446$ , 補正後  $p = .005$ ）。図より、解読力・他者受容が高い人ほど規則の使用回数が多く、推定および総回答数はほぼ一定であることが読み取れる。

一方、推定の使用回数および総回答数は、ENDCOREs を含むいずれの心理測定尺度とも有意な相関は認められなかった。また、その他の心理測定尺度（認知的欲求、多次元共感性尺度など）は、すべての推論形式と有意な相関は見られなかった。以上より、コミュニケーションのスキルが高い人ほど規則的推論を多く使用し、推定は個人特性に関わらず一定程度使用されることが示された。

表 6.5.3 個人特性と各推定形式および総回答数の相関

個人特性	推定	規則	総回答数
ENDCOREs : 他者受容	-0.011 (p=0.946, p'=0.946)	0.413* (p=0.010, p'=0.050)	0.230 (p=0.165, p'=0.330)
ENDCOREs : 解読力	0.175 (p=0.294, p'=0.601)	0.446* (p=0.005, p'=0.050)	0.378 (p=0.019, p'=0.109)
認知的欲求	-0.027 (p=0.872, p'=0.946)	0.092 (p=0.581, p'=0.581)	0.131 (p=0.434, p'=0.482)
MES : 想像性	-0.213 (p=0.199, p'=0.601)	-0.148 (p=0.375, p'=0.421)	-0.138 (p=0.408, p'=0.482)
MES : 他者指向的反応	0.213 (p=0.198, p'=0.601)	0.228 (p=0.169, p'=0.281)	0.341 (p=0.036, p'=0.120)
MES : 自己指向的反応	0.114 (p=0.495, p'=0.826)	-0.147 (p=0.379, p'=0.421)	-0.050 (p=0.766, p'=0.766)
MES : 被影響性	-0.278 (p=0.091, p'=0.601)	-0.257 (p=0.119, p'=0.281)	-0.371 (p=0.022, p'=0.109)
MES : 視点取得	0.172 (p=0.301, p'=0.601)	0.197 (p=0.237, p'=0.338)	0.303 (p=0.064, p'=0.161)
完結欲求 : 決断性	-0.018 (p=0.912, p'=0.946)	0.238 (p=0.151, p'=0.281)	0.137 (p=0.410, p'=0.482)
完結欲求 : 予測可能性選好	-0.038 (p=0.823, p'=0.946)	-0.249 (p=0.132, p'=0.281)	-0.185 (p=0.267, p'=0.445)

無記号 : 相関係数, p は p 値, p' は FDR 補正後の p 値

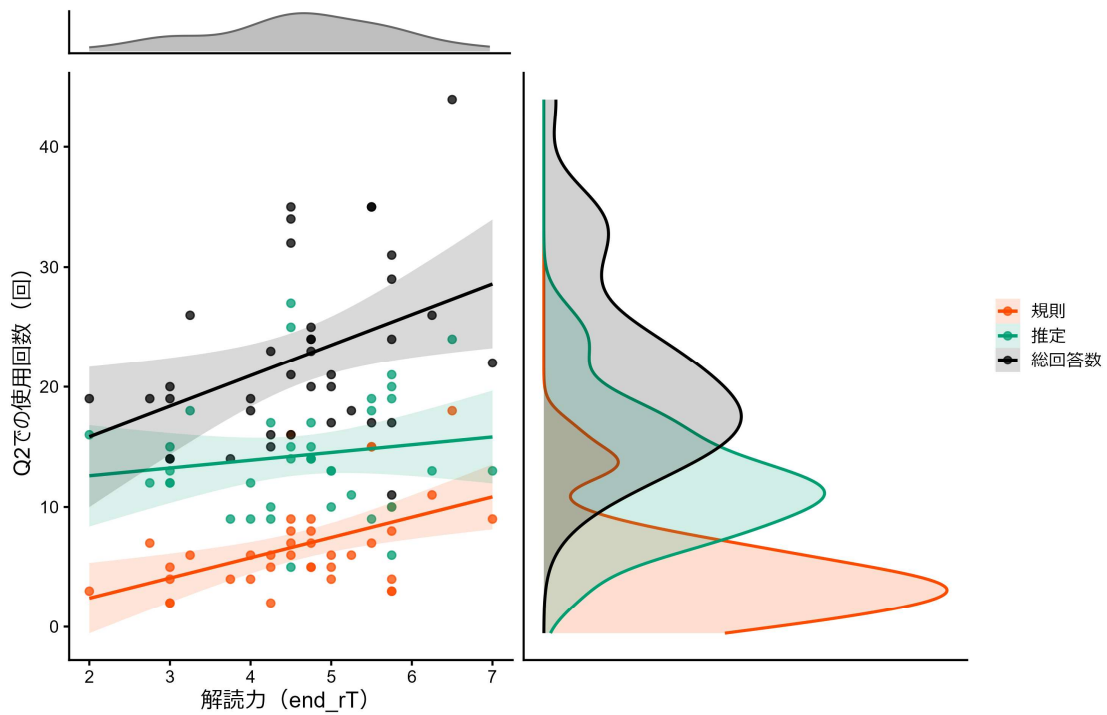


図 6.5.1 解読力と各推論形式および総回答数の相関  
 プロットは参加者

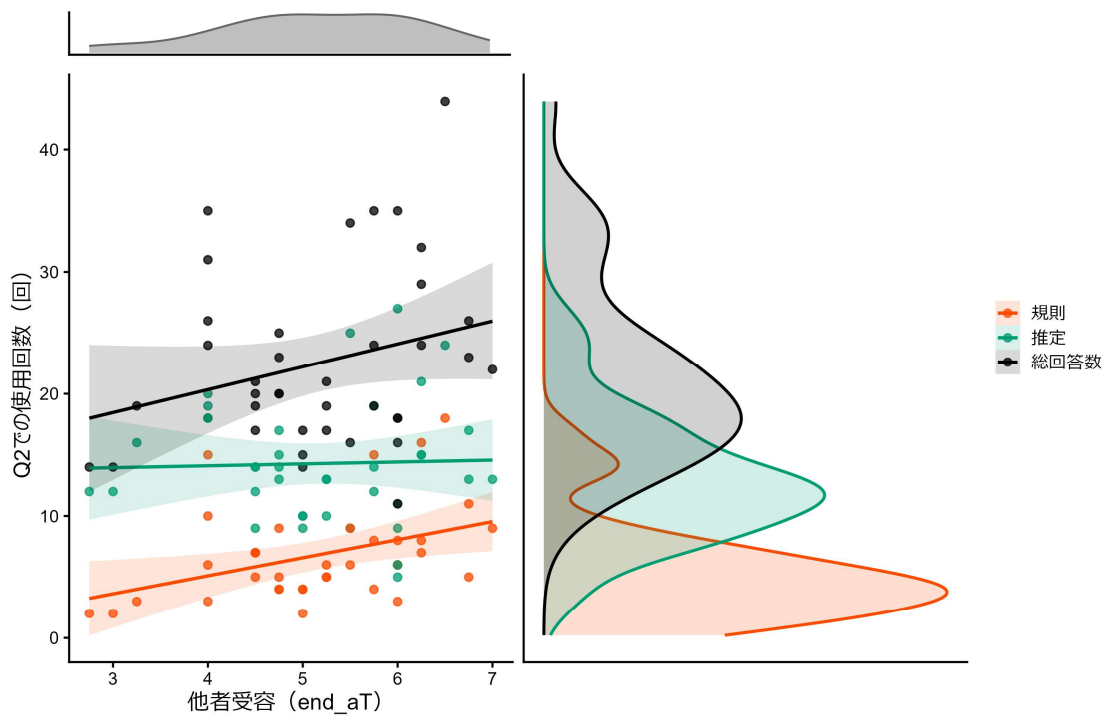


図 6.5.2 他者受容と各推論形式および総回答数の相関  
 プロットは参加者

### 6.5.3 個人特性が与える推定形式使用比率への影響

個人特性が、推論形式の使用比率ではどのような傾向があるかを探索的に検証するため、各課題の推論形式使用比率でスピアマンの順位相関を実施した。多重比較には Benjamini-Hochberg 法を用いた FDR 補正を行った。分析対象は、使用回数と同様とした。

推定形式ごとのスピアマンの順位相関の結果を表 6.5.4, 6.5.4 に示す。また、各推論形式および総回答数と有意な相関が出た個人特性（解読力および他者受容：ENDCOREs）の相関図を図 6.5.3, 6.5.4 に示す。

**表 6.5.4 個人特性と「推定」使用比率の相関**

個人特性	相関係数	p:補正前	p:補正後	star
解読力	-0.456	0.004	0.029	*
他者受容	-0.439	0.006	0.029	*
認知的欲求	-0.272	0.098	0.196	
被影響性	0.184	0.270	0.386	
他者指向的反応	-0.042	0.801	0.801	
想像性	-0.114	0.497	0.552	
視点取得	-0.127	0.447	0.552	
自己指向的反応	0.334	0.041	0.135	
決断性	-0.292	0.075	0.188	
予測可能性に対する選好	0.233	0.159	0.265	

**表 6.5.5 個人特性と「規則」使用比率の相関**

個人特性	相関係数	p:補正前	p:補正後	star
解読力	0.408	0.011	0.055	†
他者受容	0.411	0.010	0.055	†
認知的欲求	0.097	0.564	0.770	
被影響性	-0.148	0.375	0.625	
他者指向的反応	-0.034	0.841	0.841	
想像性	-0.073	0.665	0.770	
視点取得	0.066	0.693	0.770	
自己指向的反応	-0.203	0.223	0.445	
決断性	0.265	0.108	0.360	
予測可能性に対する選好	-0.216	0.193	0.445	

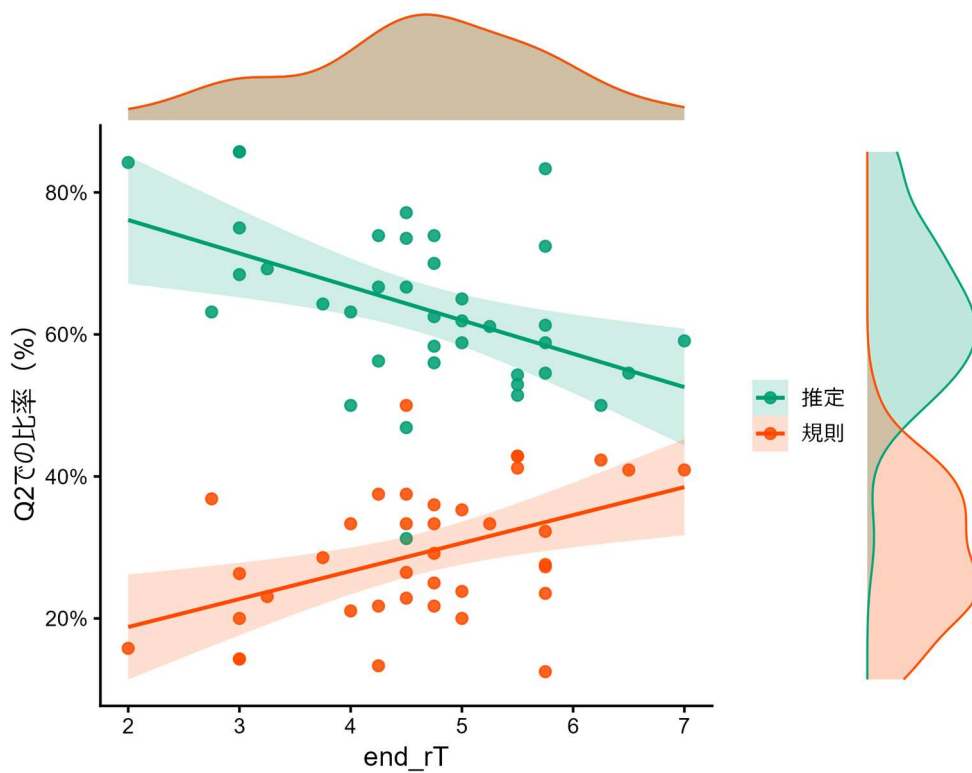


図 6.5.3 解読力 (end\_rT) と各推論形式使用比率の相関  
 プロットは参加者

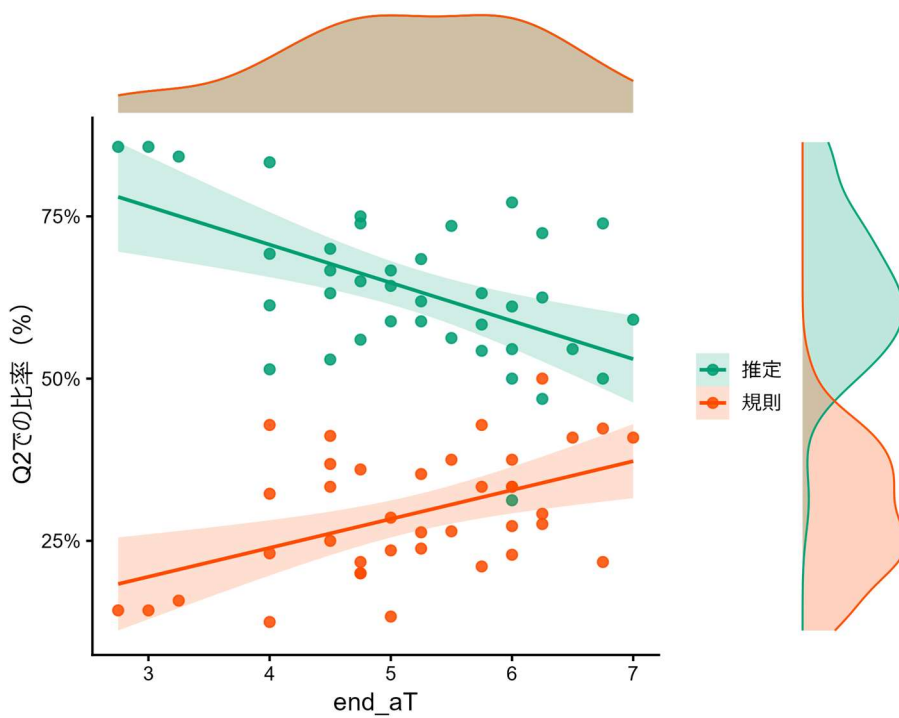


図 6.5.4 他者受容 (end\_aT) と各推論形式使用比率の相関  
 プロットは参加者

ENDCOREs の解読力および他者受容と、推定の使用比率との間に有意な負の相関がみられた（解読力： $\rho = -0.456$ , 補正後  $p < .001$ , 他者受容： $\rho = -0.439$ , 補正後  $p = .006$ ）。一方、規則の使用比率とは有意な正の相関がみられた（解読力： $\rho = 0.408$ , 補正後  $p = .011$ , 他者受容： $\rho = 0.411$ , 補正後  $p = .010$ ）。

散布図より、解読力・他者受容が高い人ほど推定の使用比率が低く、規則の使用比率が高いことが読み取れる。その他の心理測定尺度（Big Five, 認知的欲求, 多次元共感性尺度など）は、いずれの推論形式とも有意な相関は認められなかった。

#### 6.5.4 使用回数と使用比率の対比

使用回数では、推定は個人特性と有意な相関を示さず（解読力： $p = .029$ , 他者受容： $p = .029$ , いずれも補正後非有意）、規則のみが正の相関を示した。一方、使用比率では推定が負の相関、規則が正の相関を示し、両推論形式ともに有意な関連が認められた。

この違いは、二つの要因が考えられる。一つ目は推定の絶対的な使用回数は個人特性に関わらずほぼ一定である。二つ目は、規則の使用回数が増加することで総回答数が増えた。その結果、推定の相対的な比率が低下することを反映している。したがって、使用比率における推定の負の相関は、推定自体の減少ではなく、規則の増加による相対的な変化である可能性を示唆する。

## 6.6 課題間での推論形式の変化

### 6.6.1 各課題の推論形式使用比率の探索

本実験の課題では、直感回答課題（Q1）では直感的に考え、解釈選択課題（Q3）では選択を観察している。そのため、実際の発話解釈での「直ぐに解釈して、選択している」別々のプロセスを独立に観察することが可能だと考える。近い現象で推論の方略がどのようになっているかを探索的に検討する。始めに各課題の推論使用傾向を見るために、各課題の使用比率の推移（図 6.6.1）を示す。本分析では使用比率を用いる理由は、生成数では Q2 では多数回答させているため単純に比較することが困難だからだ。また、各課題の推論形式使用比率の差分図を図 6.2.2 示す。この差分図では、0 を基準とし 95%信頼区間が 0 をまたがない場合は、使用比率が参加者全体で一貫していると解釈でき、0 をまたぐ場合は使用

比率の方向は示されるものの断言は避け、傾向を見る。そして、課題別の推論形式使用回数と比率を表 6.2.1 示す。

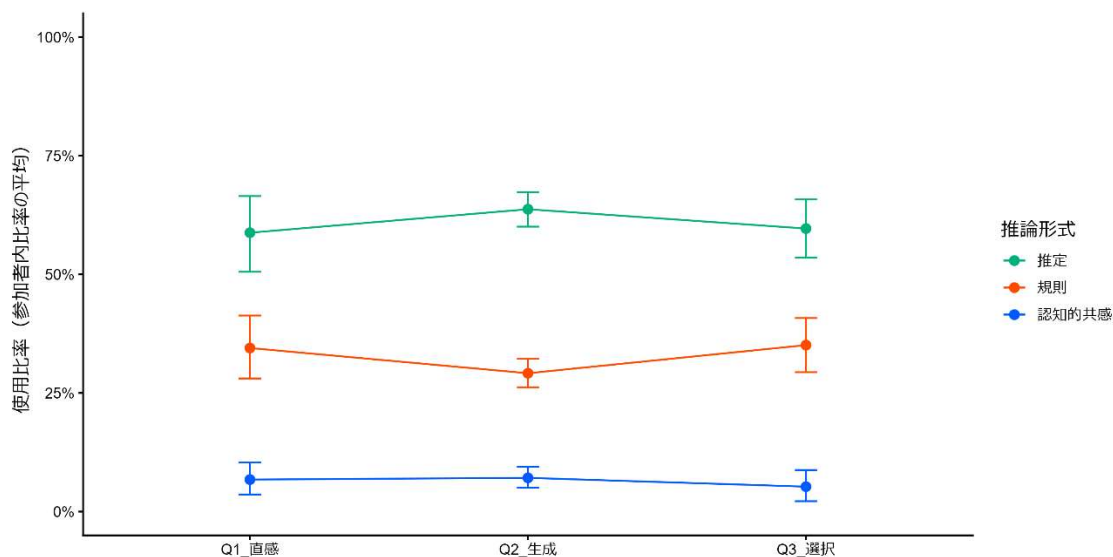


図 6.6.1 各課題の推論形式使用比率の推移

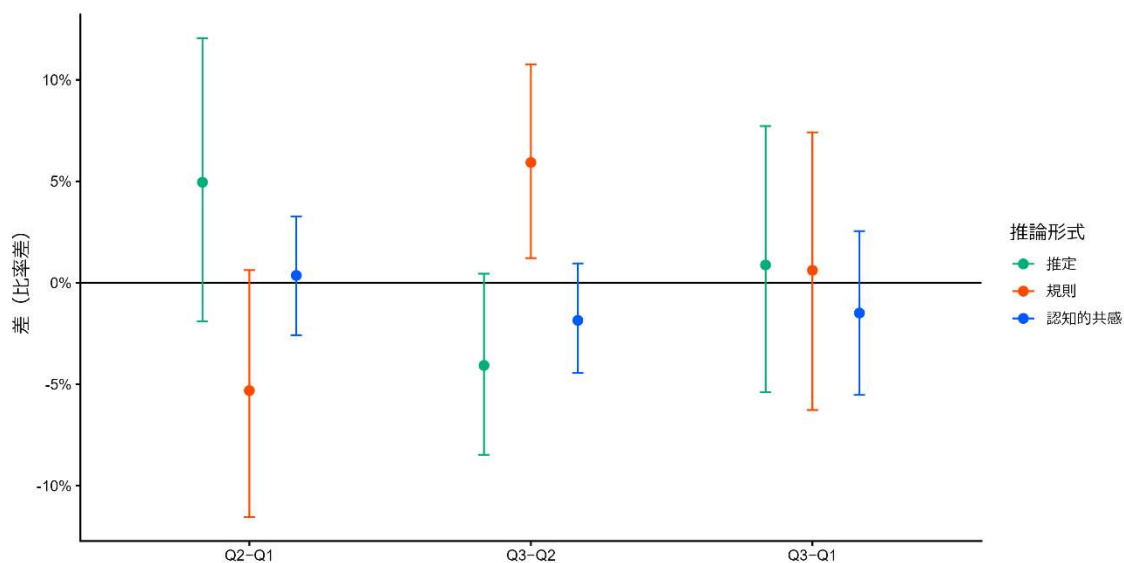


図 6.6.2 課題の推論形式使用比率の差

注：各点は平均比率，エラーバーは反復測定に合わせて参加者単位でブートストラップにより推定した 95%信頼区間である。ブートストラップとは、データを復元抽出（一度抽出し、母体に戻したからまた引ける方法）で何回も作りなおし（今回は 5000 回）、データから推定した結果のばらつきを算出する方法である。

表 6.6.1 各課題の推論形式使用比率

課題	推論形式	比率
Q1	推定	0.582
Q1	規則	0.343
Q1	認知的共感	0.075
Q2	推定	0.634
Q2	規則	0.303
Q2	認知的共感	0.063
Q3	推定	0.596
Q3	規則	0.351
Q3	認知的共感	0.053

図 6.2.2 より，熟慮回答課題では推定の使用率が少し上昇(63.4%)し，選択では減少(59.6%)傾向がある．一方で，規則は熟慮回答課題で少し使用率が低下(30.3%)したが，選択では少し上昇(35.1%)していることが読み取れる．また，図 6.2.3 より，解釈選択では，多数回答に比べて規則に分類される解釈が相対的に増える可能性を示唆する．

以上をまとめると，いずれの課題においても推定が最も高い使用比率を示しており，本研究で実施した課題の条件では，推定(アブダクション)を使った推論がされやすい傾向が示唆される．次節では，本研究の分析単位である 6 刺激から推論形式使用比率を観察する．

### 6.6.2 各刺激の推論形式使用比率の探索

次は分析単位である各刺激で推論形式の使用比率の傾向を確認する．各刺激の推論形式使用回数比率を表 6.2.2 示す．各刺激から，参加者から導出した推論形式使用比率の推移を図 6.6.3 に示す．また，各刺激の推論形式使用比率の差分図を図 6.6.4 示す．

表 6.6.2 各刺激の各推論形式使用比率

課題	刺激	推論形式	比率
直感回答	A	推定	0.57
直感回答	A	規則	0.35
直感回答	A	認知的共感	0.08
直感回答	B	推定	0.45
直感回答	B	規則	0.55
直感回答	B	認知的共感	0
直感回答	C	推定	0.73
直感回答	C	規則	0.08
直感回答	C	認知的共感	0.19
直感回答	D	推定	0.47
直感回答	D	規則	0.53
直感回答	D	認知的共感	0
直感回答	E	推定	0.56
直感回答	E	規則	0.41
直感回答	E	認知的共感	0.04
直感回答	F	推定	0.72
直感回答	F	規則	0.16
直感回答	F	認知的共感	0.13
多数回答	A	推定	0.63
多数回答	A	規則	0.32
多数回答	A	認知的共感	0.05
多数回答	B	推定	0.62
多数回答	B	規則	0.34
多数回答	B	認知的共感	0.05
多数回答	C	推定	0.69
多数回答	C	規則	0.17
多数回答	C	認知的共感	0.14
多数回答	D	推定	0.52
多数回答	D	規則	0.45
多数回答	D	認知的共感	0.03
多数回答	E	推定	0.59
多数回答	E	規則	0.39

多数回答	E	認知的共感	0.02
多数回答	F	推定	0.79
多数回答	F	規則	0.10
多数回答	F	認知的共感	0.11
解釈選択	A	推定	0.58
解釈選択	A	規則	0.32
解釈選択	A	認知的共感	0.11
解釈選択	B	推定	0.45
解釈選択	B	規則	0.53
解釈選択	B	認知的共感	0.03
解釈選択	C	推定	0.74
解釈選択	C	規則	0.18
解釈選択	C	認知的共感	0.08
解釈選択	D	推定	0.45
解釈選択	D	規則	0.53
解釈選択	D	認知的共感	0.03
解釈選択	E	推定	0.47
解釈選択	E	規則	0.50
解釈選択	E	認知的共感	0.03
解釈選択	F	推定	0.90
解釈選択	F	規則	0.05
解釈選択	F	認知的共感	0.05

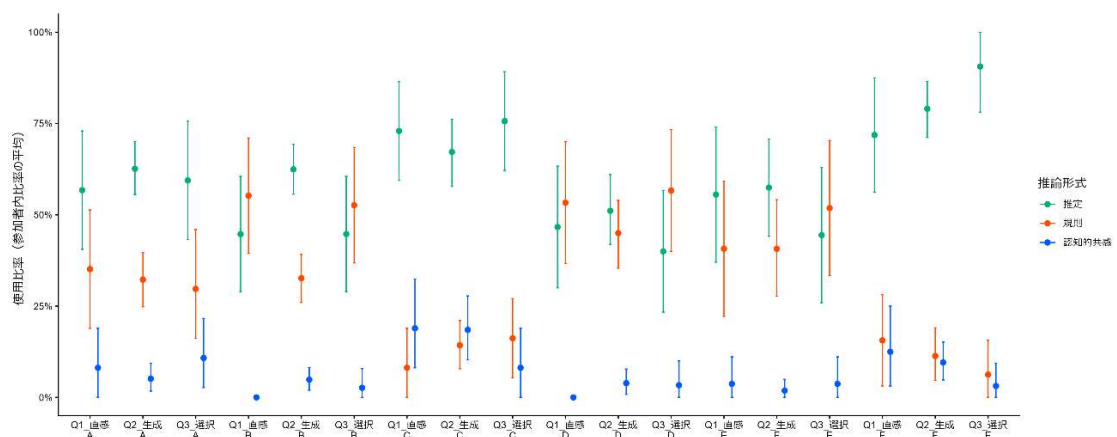


図 6.6.3 各刺激の推論形式使用比率の推移  
 各点は平均比率，エラーバーは反復測定に合わせて参加者単位で  
 ブートストラップ（5000回）により推定した95%信頼区間

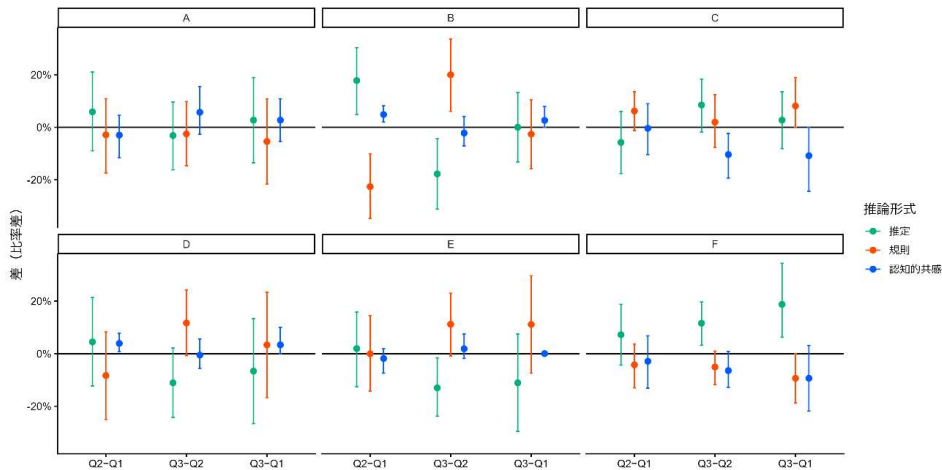


図 6.6.4 各課題の推論形式使用比率の差

各点は平均比率，エラーバーは反復測定に合わせて参加者単位でブートストラップ（5000回）により推定した95%信頼区間

図 6.6.3 より，推定の推論形式使用比率には 4 種類の傾向を確認した．1 つ目は，直感回答と解釈選択は同じくらいの使用比率で多数回答が上昇している山型の傾向 (AB)．2 つ目は，1 つ目と直感回答と多数回答は同じだが，選択では直感回答より減少している傾向(DE)．3 つ目は 1 つ目とは真逆の傾向(C)．4 つ目は，課題が進むにつれて使用比率が上昇する傾向(F)．規則では，3 種類の傾向を確認した．1 つ目は課題が進むにつれて使用比率が上昇する傾向(AF)．2 つ目は，多数回答が減少し，直感回答と解釈選択は同じ使用比率である下山型の傾向(BD)．3 つ目は課題が進むにつれて使用比率が上昇する傾向(CE)．認知的共感は数が少なく傾向を確認できなかった．

図 6.6.4 より，3 つの傾向が出た．刺激 B では，解釈選択では，多数回答に比べて規則に分類される解釈が相対的に増える可能性を示唆する．一方で，推定に分類される解釈が多数回答から解釈選択に向けて減少した可能性も示唆する．刺激 DE の規則は刺激 B と少し同じ傾向にあることが分かる．2 つ目は，刺激 F のみ多数回答および解釈選択で推定に分類される解釈が相対的に増える可能性を示唆する．

以上をまとめると，いずれの課題においても推定が最も高い使用比率を示している．本研究で実施した課題の条件では，刺激 F では推定を使った推論がされやすく，規則(帰納・演繹)を使った推論が困難である傾向が示唆される．他の刺激は各刺激で別の傾向を示した．次節では，参加者ごとに注目し個人差を探索する．

### 6.6.3 参加者ごとの見た推論形式使用割合

前節では、刺激ごとに推論形式使用比率がどのように変化したかを調べた。しかし、個人差を捉えるのは困難であり、個人レベルでの推論形式の差を確認するために、参加者ごとの Q1・Q3 の推論形式一致率を算出した。比較可能数が 5・6 のみ抽出した（表 6.6.3）。全参加者の結果は付録 G に示す。一致率の計算は以下の通りである。

$$\text{一致率} = 1 - \frac{\text{Q1・Q3 の不一致数}}{\text{Q1・Q3 がどちらも試行できた数}}$$

表 6.6.3 各参加者の Q1・Q3 推論形式一致率

参加者 ID	比較可能数	Q1・3 不一致数	一致率	推定 一致数	規則 一致数	認知的共感 一致数
101	6	1	0.83	3	2	0
104	6	2	0.67	2	2	0
105	5	1	0.80	3	1	0
106	6	0	1.00	5	1	0
110	6	2	0.67	2	2	0
111	6	0	1.00	4	2	0
112	5	2	0.60	2	1	0
113	5	0	1.00	4	1	0
114	6	2	0.67	2	1	1
116	6	1	0.83	3	2	0
117	5	1	0.80	3	1	0
118	6	2	0.67	1	2	1
119	6	2	0.67	1	3	0
121	5	2	0.60	2	1	0
122	5	3	0.40	2	0	0
123	6	1	0.83	2	2	1
124	6	1	0.83	3	2	0
125	5	3	0.40	2	0	0
126	5	2	0.60	3	0	0
128	6	1	0.83	2	3	0
129	6	2	0.67	3	1	0

131	6	3	0.50	1	1	1
132	6	3	0.50	3	0	0
133	6	1	0.83	5	0	0
134	6	0	1.00	4	2	0
135	6	0	1.00	4	2	0
137	6	3	0.50	0	3	0
138	5	0	1.00	3	2	0
140	5	0	1.00	4	1	0
142	6	1	0.83	2	3	0
143	6	4	0.33	1	0	1

比較可能数は、Q1・Q3のどちらも回答できているもののみを対象としているため、Q1でタイムアウトした人ほど数値は低くなる。本観察では、比較可能数5・6(N = 31)抽出した。全参加者は付録に示す。参加者ごとの推論形一致率を見ると一致率を見ると、Q1とQ3は8割の参加者は6刺激の内、50%は一致しており、最も一致している推論形式は「推定」であった。

#### 6.6.4 課題における推論形式の推移

以上の結果をまとめると、課題（直感／多数回答／選択）が変化しても、推論形式の平均的な序列は概ね維持され、いずれの課題でも推定が最も高かった。一方で、刺激によっては課題に伴う比率の変化も観察され、また個人内の一致率も一定程度高かった。以上より、推論形式の選択は課題要求のみでは説明しきれず、個人差の関与が大きい可能性が示唆される。

さらに6.4節では刺激間差の効果量（Kendall's W）が小さく、刺激差より個人差が大きいことが示唆された。これらを踏まえると、推論形式の選択は状況（課題・刺激）の影響を受けつつも、一定の安定性をもつ個人差に支えられている可能性が高い。

## 第7章 考察

6章では発話解釈において参加者が用いる推論形式が、(1)刺激の理解容易度、(2)個人特性によってどのように変化するかの結果を示した。本章では、これらの結果を踏まえ、仮説検証の結果と発話解釈における推論形式の関係について考察する。7.1節では、結果の概観について述べる。7.2節では、本研究の仮説検証の結果について述べる。7.3節では、発話解釈で用いる推論形式の関係について本研究の結果から議論する。7.4節では、本研究の限界について整理する。

### 7.1 結果の概観

本研究では、発話解釈においてどのような推論形式を使用しているのかを明らかにすることを目的とし、以下の2つの仮説を検証した。

仮説1：理解容易度が低いほど、アブダクションの使用頻度が増加する。

仮説2：コミュニケーションと思考特性に関する個人特性は、推論形式の使用頻度と関連する

実験刺激の操作チェックの結果は、2つの傾向を示した。反応時間では、理解容易度として想定していた3群構造（高：AB，中：CD，低：EF）で有意な差を確認した。一方、タイムアウト率では2群構造（高：ABC，低：DEF）となった。特に、刺激Cは「中」と想定していたが「高」群と同等の低いタイムアウト率（2.6%）を示した。この結果より理解容易度には異なる2つの側面を捉えていることを示唆する。反応時間は連続変数であり、処理負荷の段階的な変化を捉えることができる。一方、タイムアウトは「制限時間内に回答できない」という質的な困難さを示す指標である。刺激DとEの間には、反応時間に有意差が認められた（ $p = 0.03$ ）が、タイムアウト率には有意差が認められなかった。これは、刺激Dが解釈に時間はかかるが、制限時間以内に何とか回答が可能である難しさがある。一方、刺激EはDよりも解釈に時間がかかるが、「回答が可能であるか」という質的な難しさは刺激Dと差がないと考えられる。以降の考察では、この結果を前提に解釈を進める。

次に Think Aloud 法による推論プロセスを分析した結果、以下のことが明らかとなった。推論形式の同定として、多数回答課題における分析の結果、推定（以降、アブダクション）、規則（以降、演繹）、認知的共感の3つの推論形式が同定された。例として最も回答数が多かった刺激Bでは推定が主要な推論形式とし

て使用され（回答数：107，比率：61.5%），規則が補完的に使用された（回答数：59，比率：33.9%）。一方，認知的共感はほとんど使用されなかった（回答数：8，比率：4.6%）。

推論形式使用回数は，多数回答課題における推定の使用回数は刺激間で有意差が確認された。事後比較では，刺激 F と刺激 A の間で有意差が見られた。規則の使用回数も刺激間で有意差が見られ，事後比較では刺激 DA, DB, EA, EB, EC 間で有意差が確認された。認知的共感の使用回数は刺激に差があるかのみ検証し，刺激間で有意差が見られた。

推論形式使用比率の差について，多数回答課題におけるアブダクション使用比率は全体として最も高く（63.4%），刺激間で有意差が確認された。事後比較では，刺激 FA, FB, FD 間で有意差が見られ，刺激 F で最も比率が高く（78.6%），刺激 A で最も比率が低かった（52.5%）。演繹使用比率は刺激間で有意差が確認され，事後比較では刺激 FD, FE, AD, AE, BD, BE 間で有意差が見られた。認知的共感は刺激間に差があるかのみ分析し，有意差が確認された。

個人特性と推論形式の関係について，個人特性であるコミュニケーション・スキルと推論形式使用との相関分析では，演繹と解読力および他者受容に有意な相関が見られた。一方，アブダクションおよび総回答数とは相関が見られなかった。また，コミュニケーション・スキルと推論形式使用比率では，アブダクションと解読力および他者受容に有意な相関が見られ，演繹とも有意傾向の相関が見られた。

課題の性質（直感・多数・選択）が変わっても，推論形式を使用が変化するか探索的に分析した。結果，推論形式の使用傾向は概ね維持され，使用パターンは大きくは変化しなかった。この結果は，推論形式の選択が課題要求そのものよりも，個人差が影響した可能性を示唆する。また，参加者単位でも推論形式が一定程度一致する傾向が確認され，「その人らしさ」が反映されていると示唆する。次節では，仮説 1 の検証結果と解釈について考察を述べる。

## 7.2 理解容易度と推論形式の関係

### 7.2.1 仮説の検証結果

仮説 1 「理解容易度が低いほど，アブダクション(推定)の使用が増える」について，本研究では支持されなかった。具体的には，推定使用回数は刺激間で有意差が確認された ( $\chi^2(5)=12.40, p=.030$ ) が，理解容易度の代表値に基づく比較

(高：AB, 中：CD, 低：EF) では有意差が見られなかった。また、アブダクション使用回数の平均値においても、同傾向を示した (A：2.74, B：2.82, C：2.16, D：2.03, E：1.95, F：2.61)。以上の結果より、想定した理解容易度の順序 (高：AB, 中：CD, 低：EF) に対応するアブダクションの増加は見られなかった。

### 7.2.2 結果から考えられる発見

この仮説 1 の非支持という結果は、次の発見ができたのではないかと考えられる。本研究では当初、理解が難しい発話ほど、既存の知識や規則に当てはめることが難しく、アブダクションに頼る必要があると仮定していた。この見方は、アブダクションを「演繹ができない場合に使う手段」として想定していた。

しかし、本研究の結果はこの想定を否定する。アブダクションは、理解困難な状況だけで使われる「代替手段」ではなく、理解容易度に関わらず常に使用される推論形式であると考えられる。実際、すべての刺激において推定が最も多く使用され、演繹や認知的共感を大きく上回った。

この発見は Sperber and Wilson (1986/1995) の関連性理論の記述と整合的であり、推論形式の内実を経験的に示唆するものだと考えられる。関連性理論では、発話解釈における推論メカニズムは、主に演繹であると想定する一方、非論証的でもあると記述され、推論形式が明確ではなかった (Sperber & Wilson, 1986/1995; 吉村, 2016)。これに対し、本研究の実際の発話解釈における認知プロセスでは (1) 複数の推論形式が使用されていること、(2) その中でアブダクションが主に使用されていること、(3) 推論形式の方略には個人差があること、を実証的に示した。この知見は、実際の発話解釈における推論メカニズムが多様であることを明らかにした点で、関連性理論への経験的貢献であると考えられる。

### 7.2.3 仮説 1 非支持の要因

仮説 1 が支持されなかった要因として、理解容易度という概念は単純すぎた可能性がある。予備調査および本実験の結果から、理解容易度は単一の「易しい-難しい」次元では捉えきれない、多次元的な構造を持つことが示唆された。本研究で理解容易度を「中」に分類した刺激 C・D は、本実験での推論パターンが大きく異なった。具体的には、推定使用回数は刺激 C：M=2.16, 刺激 D：M=2.03, 規則使用回数\_刺激 C：M=0.53, 刺激 D：M=1.74, と同じカテゴリーで異なる性質が示された。また、実験刺激の操作チェックでは、反応時間では 3 群構造

(高：AB, 中：CD, 低：EF) が, タイムアウト率では 2 群構造 (高：ABC, 低：DEF) となった. これは, 理解容易度が想定通りに操作できていない可能性を示唆する.

また、解釈の生成数の観点から考察すると、本研究の刺激は、解釈を導出しやすくするために、Grice (1975) の協調の原理における公準 (様態・関係) からの逸脱を意図的に組み込んでいた. この設計意図に基づけば、公準逸脱が大きいほど不足情報を補う探索的推論 (アブダクション) が増加する、という予測が立つ. しかし結果は、理解容易度の想定順序に対応した単調増加としては観察されなかった.

この不一致は、公準からの逸脱が推意を考えさせる「きっかけ」にはなる一方で、それだけで解釈候補が必ず増えるとは限らないということを示唆している. 例として、様態の公準の逸脱は、解釈を促しつつも、社会規範や経験的規則が利用可能であれば推論が早期に収束しやすい可能性がある. また関係の公準の逸脱が強い場合でも、関連性理論の理解手続きに従えば、解釈は最小の労力で進み、期待される関連性が満たされた時点で探索は停止するため、候補生成が際限なく増えるとは限らない. 加えて、本研究の生成数は Think Aloud により外化された指標であるため、関連づけが困難な状況では探索が増える以前に言語化が難しくなり、生成数として表れにくい可能性もある. 以上より、本研究で作成した刺激群は「理解容易度」という一軸に等質に並ぶのではなく、語彙の難しさや会話の対象といった複数次元の組み合わせとして捉える必要があり、そのことが仮説 1 の不支持につながった可能性が示唆される.

## 7.3 個人特性と推論形式の関係

### 7.3.1 仮説 2 の検証結果

仮説 2「コミュニケーションと思考に関する個人特性は、推論形式の使用頻度と関連する」については、部分的に支持された. 具体的には、コミュニケーション技能 (ENDCOREs: 解読力・他者受容) が推論形式の使用回数および使用比率と関連することが示された. 使用回数では、演繹が正の相関傾向 (解読力:  $\rho = .413$ , 補正後  $p = .050$ , 他者受容:  $\rho = .446$ , 補正後  $p = .050$ ) を示した. アブダクションおよび総回答数は有意な相関を示さなかった.

使用比率では、アブダクションが負の相関 (解読力:  $\rho = -.456$ , 補正後  $p = .029$ , 他者受容:  $\rho = -.439$ , 補正後  $p = .029$ ) を示した. また、演繹が正の相関傾向

(解読力： $\rho = .408$ , 補正後  $p = .055$ , 他者受容： $\rho = .411$ , 補正後  $p = .055$ ) を示した。以上の結果は、コミュニケーション技能が推論形式の使い方に影響するが、その影響はアブダクションの“増減”ではなく、演繹の“追加”として現れる可能性を示唆する (図 7.3.1 に再掲)。

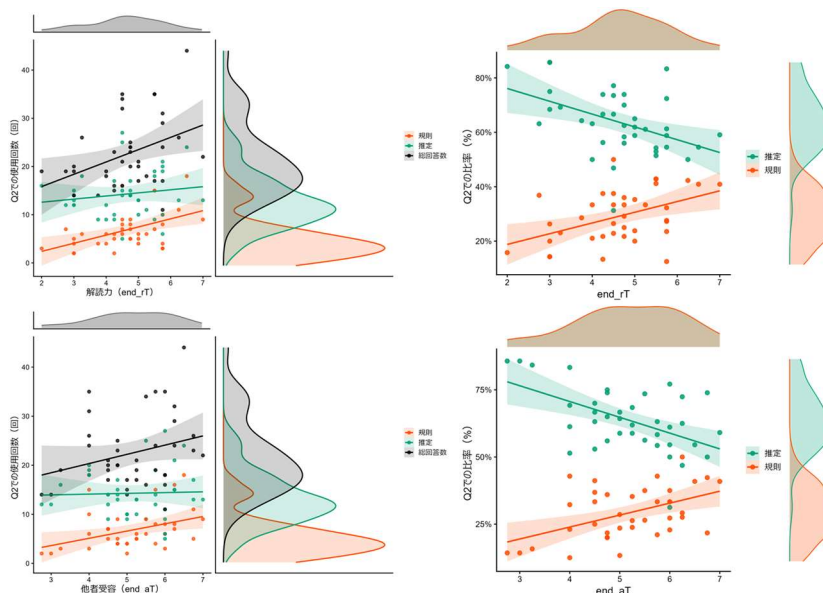


図 7.3.1 推論形式使用回数・比率と解読力・他者受容の相関

左上：使用回数×解読力，右上：使用比率×解読力

左下：使用回数×他者受容，右下：使用比率×他者受容

### 7.3.2 演繹追加方略

この結果は「高スキル者は演繹を使い、低スキル者はアブダクションを使う」といった、演繹とアブダクションのどちらか一方しか使わず、スキルによって「アブダクションから演繹へ入れ替わる」と単純なものでは無い。むしろ、アブダクションは解釈の難しさやコミュニケーション・スキルに関わらず広く使用されており、コミュニケーション・スキルの高い者は演繹を追加的に使用するという拡張方略として捉える方が適切だと考える。結果より、アブダクションの使用回数は解読力・他者受容が関連せず、演繹の使用回数のみが正の相関を示した。これは「アブダクションが減少した」のではなく、「アブダクションと同時に演繹も使用されている」可能性を示唆する。

この解釈は、解読力・他者受容の心理学的な意味とも整合する。解読力とは相手の表情・声・言い方・文脈など複数の手がかりから「相手がどのような気持ちか」を読み取る力である。他者受容は相手の立場・感情・価値観を尊重して受け

止める対人スタンスである。これらが高い人ほど、対話を「難しい」と感じにくく、認知的な余裕から解釈を探索的に行えるだろう。

この知見は、Sperber and Wilson (1986/1995)の関連性理論とも親和的である。聞き手は最も妥当な解釈を推定しつつ、必要最小限の労力で、最大の関連性である解釈を探り収束していく。本研究の知見は、Sperber and Wilson (1986/1995)が想定していた、発話解釈が演繹的な推論規則であることの実証となる可能性を示唆する。

### 7.3.3 使用回数と使用比率の対比

本研究のもう一つの示唆として、使用回数と使用比率のどちらを指標とするかで、解釈が異なる点である。各推論形式の使用比率は、総回答数で割って算出される。そのため、演繹の回数が増えると総回答数が増加し、アブダクションの回数は不変でも比率は相対的に低下するだろう。

したがって、比率の結果だけを見ると、「コミュニケーション・スキルの低い者はアブダクションに依存し、高い者は演繹へ移行する」という置換として解釈してしまう。しかし、使用回数の結果と整合的に考えると、「アブダクションの減少」ではなく、高スキルの方が演繹を付加したと解釈するべきである。仮説2の検証においては、使用回数と比率を見ることで推論形式の変化を「置換」とではなく、「付加」として解釈することを可能とする。

## 7.4 統合的理解

### 7.4.1 仮説1と仮説2の整理

以上の考察を統合すると、発話解釈における推論形式の選択は「理解容易度が低いほどアブダクションが増える」という単純なモデルでは説明できず、アブダクションは解釈の難しさやコミュニケーション・スキルの高さに関わらず作動しており、演繹は解釈の難しさに関わらずコミュニケーション・スキルの高い人ほど追加的に使用していることを示唆する。

仮説1は支持されず、理解容易度の想定順序に対応したアブダクション使用の増加は観察されなかった。これは、「難しいときだけアブダクションに頼る」という見方を棄却し、むしろアブダクションが発話解釈において常時稼働する推論である可能性を示唆する。本研究の結果でも、どの刺激でもアブダクションは最も多く、演繹や認知的共感を上回った。したがって理解容易度は、アブダクシ

ョンを増減させるのではなく、アブダクションに加えて知識・慣習（演繹）をどの程度動員できるかに影響すると考えられる。

この点を仮説2の結果と合わせると、統合像がより明確になる。ENDCOREsの解読力・他者受容は、アブダクションの使用回数とは関連しない一方、演繹の使用回数と正の関連を示した。この結果は、コミュニケーション技能がアブダクションを増やすのではなく、演繹を追加的に使えるかどうかとして示している。これにより、アブダクションは多くの参加者で共通に維持されつつ、コミュニケーション・スキルの高スキル者では知識や慣習に基づく演繹的推論が“上乘せ”する推論で発話解釈は行われていると示唆される。

この統合像は、Grice (1975) の協調の原理と整合的に解釈できる。協調の原理では、聞き手は話者が会話目的に沿って発話していると仮定し、字義的内容だけでは確定しない意図を推論によって補う。ここで、文脈と経験から「この状況なら普通はこういう意図だろう」という一般化された知識が利用可能であれば、聞き手はその規則を手がかりに意図を効率よく絞り込める。そのため、解釈は演繹優勢になりやすい。一方、状況に合う規則が見当たらない／複数の規則が競合する／規則が適用しにくい場合には、候補を探索的に立てて最も可能性のある解釈を探るアブダクションが相対的に選択される。

この違いは、吉村 (2016) が区別する「二つの推意」にも対応づけられる。既知の規則から分析的に導かれやすい演繹される推意が成立しやすい状況では演繹が有利になる。それに対し、状況に即した新しい解釈候補を立てる創作される推意が強く求められる状況では、アブダクションが中心的役割を担う。仮説1が支持されなかったのは、「理解容易度（易しい／難しい）」という単一の尺度が、実際にはこの二つの推意の条件差（規則が効く状況か／新規候補が必要な状況か）を十分に分離できていなかったため、と考えられる。たとえば同じ“中”に分類したCとDで推論パターンが大きく異なった点は、理解容易度が一軸ではなく、演繹の適用可能性といった他の側面を含むことを示唆する。

#### 7.4.2 仮説1と仮説2の統合

以上を踏まえると、本研究の統合的結論は次のようにまとめられる。発話解釈では、アブダクションが解釈の難しさやコミュニケーション・スキルの高低に関わらず広く作動する推論であり、その上に演繹（知識・慣習・経験則）が状況に応じて追加される。この「追加」の生起には、(1) その場面で規則が使えるかという状況要因と、(2) 解読力・他者受容のように対話処理の負担を下げ、解釈の

手がかりを取り込みやすくする個人差要因が関与する。したがって、推論形式の選択は「難しいからアブダクション」という単純な代替関係ではなく、コミュニケーション・スキルが高い人が上乘せで演繹でも解釈が可能であると理解するのが妥当である。

## 7.5 刺激特性の複雑性と理解容易度概念の限界

本研究で使用した実験刺激は、理解容易度評定調査（オンラインアンケート）を実施して作成したが、本実験では理解容易度評定と推論メカニズムの対応の検討では、単純な対応関係では見られなかった（6.1 節参照）。この不一致は、理解容易度が単一の「分かりやすい—（理解が）難しい」次元では捉えきれない多次的な構造を持つ可能性を示唆する。

また、本研究の結果は特定の会話刺激（A～F）に基づく結果であり、同種の効果が他の話題や相手との関係性（友人・上司・家族など）や、文化の異なる場面でも同様に成立するかは保証されない。実験刺激の設計上「曖昧さ」を意図的に含めているため、自然会話における連続的な文脈更新や相互作用の影響はほとんど取り込めていない。今後の研究では、理解容易度をより精緻に定義し、返答の意味的明確性、知識依存性、推論の深さなど複数の視点から刺激を統制する必要がある。

## 7.6 本研究の限界

上記の他に、本研究には以下の4つの限界がある。一つ目に、サンプルサイズの問題がある。本研究では  $N=38$  であり、個人差について詳細な分析が難しい。また、本研究では学生を対象としており、年齢範囲が狭い。発話解釈の推論プロセスは年齢によっても変化する可能性があるため、今後のサンプルでは中高年層や高齢者のデータを増やすことができれば、より一般化可能性が上がると思われる。

二つ目に、Think Aloud 法に基づく推論形式の分類には、口頭化という測定方法に固有の制約がある。本研究で得られた推論形式の使用率は、参加者の思考過程そのものではなく、口頭報告として外化された内容に依存する。なぜなら、実際には思考していても言語化されない候補や、言語化の負荷によって生成や比較が抑制される可能性がある。したがって、本研究の推論形式使用率は「心内で生じた推論の割合」ではなく、「報告された推論の割合」として解釈する必

要がある。

三つ目は、Think Aloud 法に基づく推論形式の分類を 1 人で行ったことである。原因として、本実験で得られた反応時間および Think Aloud 法による回答データは一部（文字起こし）を除き、すべて筆者が手作業で実施した。そのため時間的制約によりコーディング方法を教えるなどの時間が足りなかった。

四つ目に、直感課題におけるタイムアウト率は、解釈の困難さを反映する有用な指標であるものの、純粋に「意図が分からない」ことだけを表すとは限らない。回答制限時間の枠内で、実験刺激の読み取り速度、回答(言語化)にかかる時間などが影響する可能性がある。そのため、タイムアウト率を理解容易度の唯一の指標として扱うことはできず、本研究でも示されたように、アブダクション使用率との不一致を含めて多面的に解釈する必要がある。以上の限界を踏まえ、次章では本論文のまとめについて述べる。

## 第8章 おわりに

7章では、仮説1および仮説2を検証する実験で得た結果を基に発話解釈における推論形式がどのように使用されるかについて考察した。本章では、研究の概要を整理し(8.1節)、主要な結論(8.2節)および今後の課題(8.3節)について述べる。

### 8.1 研究の概要

発話解釈は、字義的意味の復号だけでは完結せず、話者が何を意図しているのかを聞き手が推論によって補う過程を含む。関連性理論(Sperber & Wilson, 1986/1995)では発話解釈における推論を想定はしているが、具体的な推論形式は明示されていない(吉村, 2016)。一方で、吉村(2016)や Hashimoto(2020)は発話解釈がアブダクションで推論されていると提唱しているが、理論的検討であり実証までは行われていない。そこで本研究の目的は、発話解釈において聞き手がどのような推論形式を用いて解釈を推論しているのかを、実証的に明らかにすることである。この目的を達成するために以下の2つのリサーチクエスチョンを設定した。

RQ1: 発話の解釈における推論形式の方略はどのようなものか

RQ2: コミュニケーションと思考に関する個人特性は、推論形式の使用頻度と関連する

若年成人38名を対象に、理解容易度を3水準(高・中・低)に分類した6種類の会話刺激を用いて実験を行った。参加者は各刺激について直感的解釈を回答し、複数の解釈候補を生成し、最後に最適解釈を選択した。解釈生成過程は思考発話法により収集し、発話プロトコルを3つの推論形式にコーディングした。推論形式は、アブダクション(文脈から意図を仮説的に推測するアブダクション)、演繹(一般知識や慣習に基づく演繹的推論)、認知的共感(話者の視点に立って意図を推測する推論)の3種類である。

本研究では2つの仮説を検証した。仮説1「理解容易度が低いほどアブダクションの使用が増える」について、アブダクション使用回数は刺激間で有意差が確認された( $\chi^2(5) = 12.40, p = .030$ )。しかし、理解容易度の代表値比較では有意差が見られず、平均値の並びも想定順序(高: AB, 中: CD, 低: EF)に沿った単調増加とはならなかった(A: 2.74, B: 2.82, C: 2.16, D: 2.03, E: 1.95, F: 2.61)。したがって、理解容易度はアブダクション使用を単純に増減させる主

要因とは言えず、仮説 1 は支持されなかった。

仮説 2「コミュニケーションと思考に関する個人特性は推論形式の使用パターンと関連する」について、コミュニケーション・スキル (ENDCOREs: 解読力・他者受容) は、使用回数においてアブダクションとは相関が見られなかった。一方、演繹使用回数に有意な正の関連を示した (解読力:  $\rho = .413$ , 補正後  $p = .050$ ; 他者受容:  $\rho = .446$ , 補正後  $p = .050$ )。また使用比率では、アブダクション使用比率が有意な負の相関 (解読力:  $\rho = -.456$ , 補正後  $p = .029$ ; 他者受容:  $\rho = -.439$ , 補正後  $p = .029$ ) を示し、演繹使用比率は有意な正の相関傾向を示した (解読力:  $\rho = .408$ , 補正後  $p = .055$ ; 他者受容:  $\rho = .411$ , 補正後  $p = .055$ )。なお、心の理論に関連する尺度も使用したが、有意な相関は認められなかった。以上より、コミュニケーション・スキルの影響がアブダクションの増減という形よりも、演繹の追加という形で現れる可能性が示唆された。したがって、仮説 2 は部分的に支持された。

加えて、課題の性質 (直感/多数回答/選択) が異なっても推論形式の使用傾向が大きく変化しないことが示され、推論形式の選択は課題要求よりも個人差に強く規定される可能性が示唆された。

## 8.2 結論

本研究の結論は、次の三点である。一つ目は、発話解釈においてアブダクションは解釈の理解容易度や課題の性質に左右されにくく、発話解釈において広く用いられる推論形式だと考えられる。二つ目は、コミュニケーション・スキルが高いほど、アブダクションの増減ではなく、演繹の追加的動員として現れる。三つ目は、推論形式の選択は刺激や課題の違いよりも個人差の影響が大きく、発話解釈の推論方略は、状況依存的ではなく、一定の安定性を持つ認知スタイルとして捉える必要を示唆する。以上より、発話解釈の推論プロセスは「アブダクションは発話解釈の難しさおよびコミュニケーション・スキルによらず最も使用される推論形式であり、コミュニケーション・スキルが高い人ほど演繹が上乗せで発話解釈を行うことが可能」である可能性を示唆する。

## 8.3 今後の課題

今後の課題は次の 4 つが考えられる。一つ目は、より大規模なサンプルでアブダクションに演繹が上乗せされるという方略の再現性を確認する必要がある。

る。加えて、理解容易度と個人特性の交互作用（例：高スキル者ほど困難刺激で演繹が増えるか）も、十分な検出力のもとで検討することが望まれる。2つ目は、本実験では日本語母語話者かつ学生を対象とした。今後は、年齢層や文化圏・言語の異なるサンプルに拡張し、本研究と同様の結果が得られるかを検証する必要がある。3つ目は、本研究で作成した刺激を多次元（意味的明確性・知識依存性）で整理する必要がある。これにより、どの要因が推論形式の使い分けに影響するか更なる知見が得られると考える。4つ目は、3にも関連することがだが本研究で作成した実験刺激では、発話相手の情報（関係性など）や経験・ルールといった知識を用いた推論が行いづらい刺激の特徴を持っていることが考えられる。

## 謝辞

本研究を進めるにあたって、終始懇切丁寧なご指導を賜りました橋本敬先生には心より感謝申し上げます。未熟な研究構想を学術的水準にまで引き上げていただきました。また、ゼミでの議論や表面的な理解に留まらず本質を捉える思考力の重要性を、辛抱強くご指導くださいました。研究生活のみならず、北陸の地での貴重な経験の機会を数多く与えていただき、充実した2年間を過ごすことができました。

黒川瞬先生には、入学後の半年間という短い期間ではございましたが、研究への向き合い方について貴重なご助言をいただきました。深く感謝申し上げます。

ゼミや研究活動においては、橋本研究室のメンバーの皆様に多大なご支援をいただきました。特に笠野 純基さんには1年間心理学実験の基礎から分析方法まで色々と議論やアドバイスを頂き、深く感謝申し上げます。また、研究室の先輩である成 太俊さん、周 豪特さん、秦 慕君さん、岩村 入吹さん、勅使河原 梨沙さん、郝 抒妍さん、同期の名嘉 琉星君、中村 都夢君、Wang Zibin 君、後輩の井崎 心斗君、本那 真一君、繪内 颯太君、朱 麗雅さん、鄧 娜さんにも心より感謝申し上げます。

最後に、本研究の実験にご協力くださいました参加者の皆様、ならびに本研究に関わってくださったすべての方々に深く御礼申し上げます。

## 参考文献

- Ali, Y. H. (2020). How might simulation-based accounts of mindreading explain pragmatic interpretation? *Phenomenology and Mind*, 19, 226–240. <https://doi.org/10.17454/pam-1914>.
- Ash, I. K., Jee, B. D., & Wiley, J. (2012). Investigating insight as sudden learning. *The Journal of Problem Solving*, 4(2), 1–27. <https://doi.org/10.7771/1932-6246.1123>.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493–501. [https://doi.org/10.1016/S1364-6613\(98\)01262-5](https://doi.org/10.1016/S1364-6613(98)01262-5).
- Gallistl, M., Kunzl, M., Gabler, S., Kanske, P., Vrticka, P., & Engert, V. (2024). Attachment and inter-individual differences in empathy, compassion, and theory of mind abilities. *Attachment & Human Development*, 26(4), 350–365. <https://doi.org/10.1080/14616734.2024.2376762>.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). Academic Press.
- Hashimoto, T. (2020). The emergent constructive approach to evolinguistics: Considering hierarchy and intention sharing in linguistic communication. *Journal of Systems Science and Systems Engineering*, 29(6), 675–696. <https://doi.org/10.1007/s11518-020-5469-x>.
- Klichowicz, A., Lippoldt, D. E., Rosner, A., & Krems, J. F. (2021). Information stored in memory affects abductive reasoning. *Psychological Research*, 85, 3119–3133. <https://doi.org/10.1007/s00426-020-01460-8>.
- Mayn, A., & Demberg, V. (2025). Sources of individual variability in a pragmatic reference game: Effects of reasoning and Theory of Mind. *Cognitive Science*, 49(1), [https://doi.org/10.31234/osf.io/3kefd\\_v1](https://doi.org/10.31234/osf.io/3kefd_v1).
- Peirce, C. S. (1934). *Collected papers of Charles Sanders Peirce: Vol. 5. Pragmatism and pragmatism* (C. Hartshorne & P. Weiss, Eds.). Harvard University Press.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *NeuroImage*, 19(4), 1835–1842. [https://doi.org/10.1016/s1053-8119\(03\)00230-1](https://doi.org/10.1016/s1053-8119(03)00230-1).
- Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal cortex. *Brain*, 132(3), 617–627. <https://doi.org/10.1093/brain/awn279>.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford, Blackwell. (Original work published 1986).
- Sperber, D. & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Oxford, Basil Blackwell. 内田聖二, 中達俊明, 宋南先, 田中圭子 (訳) (2000). 関連性理論 — 伝達と認知 —, 研究社出版.

- Zambrano Cruz, R., & Suárez de la Torre, M. (2020). Effect of theory of the mind on the pragmatic reasoning of scalar implicatures. *Tesis Psicológica*, 16(1), 42–53. <https://doi.org/10.37511/tesis.v16n1a2>.
- Żelechowska, D., Żyluk, N., & Urbański, M. (2020). Find out a new method to study abductive reasoning in empirical research. *International Journal of Qualitative Methods*, 19, 1–11. <https://doi.org/10.1177/1609406920909674>.
- Zu, D., & Xia, S. (2012). Research on interpretation process of bridging reference in English discourse with think aloud protocol. *English Linguistics Research*, 1(2), 119–137. <https://doi.org/10.5430/elr.v1n2p119>.
- 秋元頼孝, 邑本俊亮. (2007). 認知的負荷がアイロニーの判断に及ぼす影響—アイロニーらしさの知覚は自己中心的視点から生じるのか—. *認知科学*, 14(3), 292-302. <https://doi.org/10.11225/jcss.14.292>.
- 神山貴弥, 藤原武弘. (1991). 認知欲求尺度に関する基礎的研究. *社会心理学研究*, 6(3), 184-192. <https://doi.org/10.14966/jssp.KJ00003725148>.
- 箕輪朗 (2025). 関連性理論に基づく意図解釈の実証的検討：認知効果・心的労力・意図解釈の関係性[修士論文, 北陸先端科学技術大学院大学先端科学技術研究科修士論文]. Japan Advanced Institute of Science and Technology Academic Repository. <https://hdl.handle.net/10119/19745>.
- 水野太貴 (2024). きょう, ゴリラをうえたよ：愉快で深いこどものいいまぢがい集, 株式会社角川.
- 鈴木公基, 桜井茂男. (2003). 認知的完結欲求尺度の作成と信頼性・妥当性の検討. *心理学研究*, 74(3), 270-275. <https://doi.org/10.4992/jjpsy.74.270>.
- 鈴木有美, 木野和代. (2008). 多次元共感性尺度 (MES) の作成 自己指向・他者指向の弁別に焦点を当てて. *教育心理学研究*, 56(4), 487-497. [https://doi.org/10.5926/jjep1953.56.4\\_487](https://doi.org/10.5926/jjep1953.56.4_487).
- 藤本学, 大坊郁夫. (2007). コミュニケーション・スキルに関する諸因子の階層構造への統合の試み. *パーソナリティ研究*, 15(3), 347-361. <https://doi.org/10.2132/personality.15.347>.
- 並川努, 谷伊織, 脇田貴文, 熊谷龍一, 中根愛, 野口裕之. (2012). Big Five 尺度短縮版の開発と信頼性と妥当性の検討. *心理学研究*, 83(2), 91-99. <https://doi.org/10.4992/jjpsy.83.91>.
- 若林明雄, 東條吉邦. (2004). 自閉症スペクトラム指数 (AQ) 日本語版の標準化 高機能臨床群と健常成人による検討. *心理学研究*, 75(1), 78-84. <https://doi.org/10.4992/jjpsy.75.78>.
- 米盛 裕二. (2007). アブダクション：仮説と発見の論理. 勁草書房.
- 吉村あき子 (2016). 演繹される推意と創作される推意. *Papers from the Thirty-*

Third Conference and from the English International Spring Forum of The English Linguistic Society of Japan, 33, pp. 209-215.

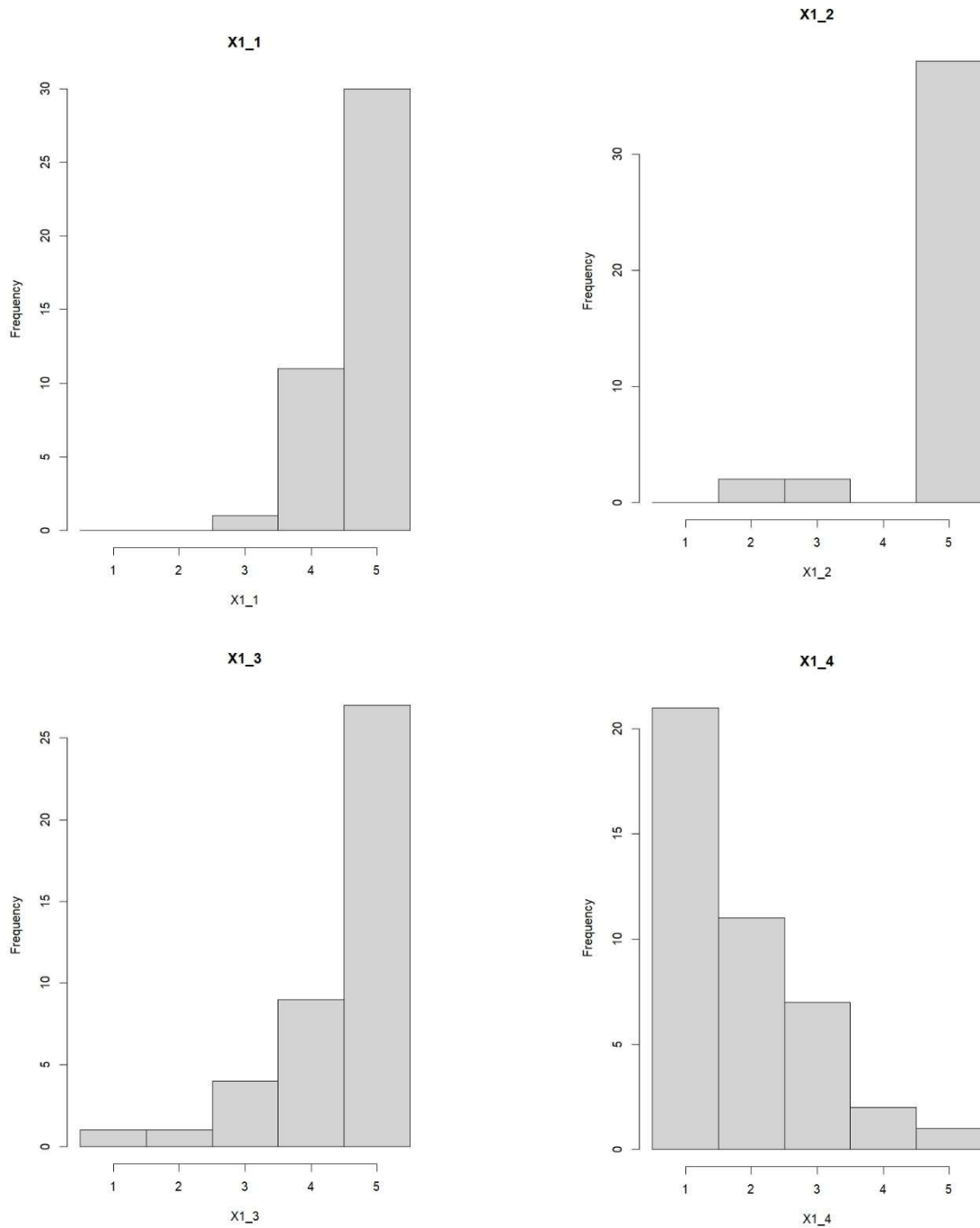
## 付録 A：理解容易度調査（3章）のために作成した会話テキスト

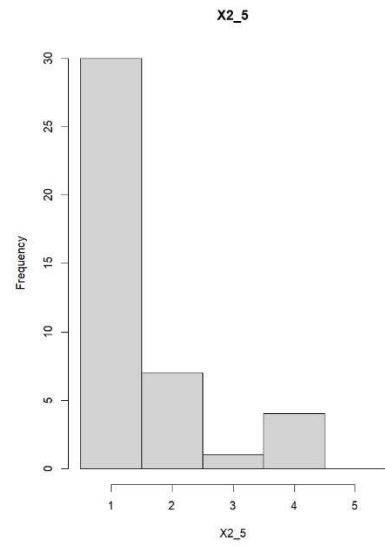
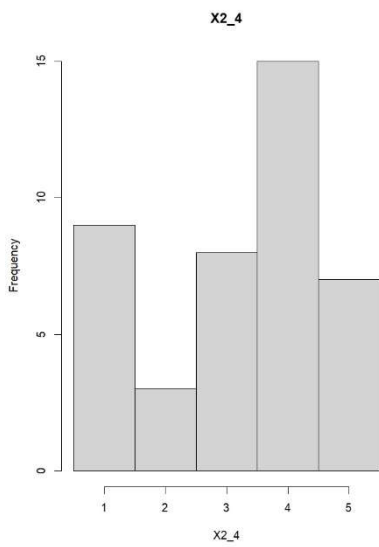
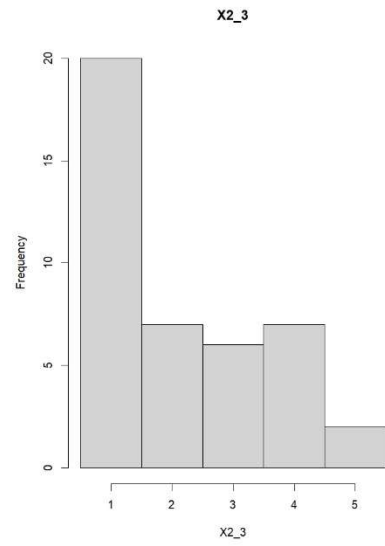
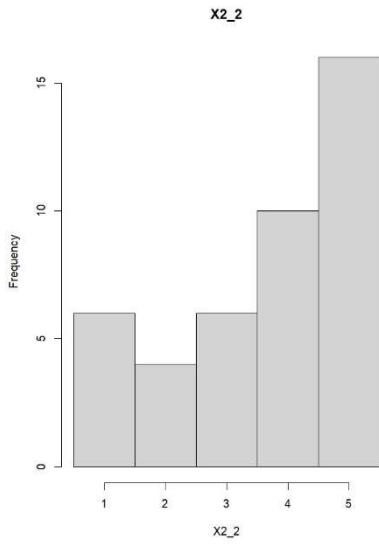
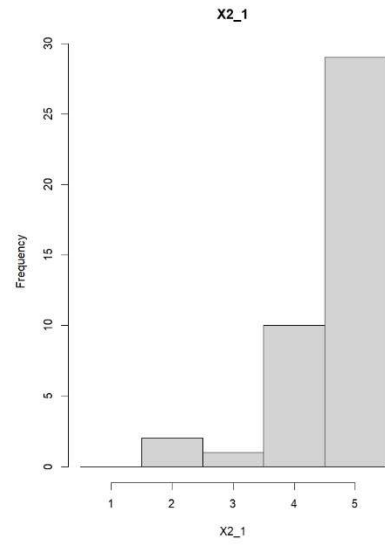
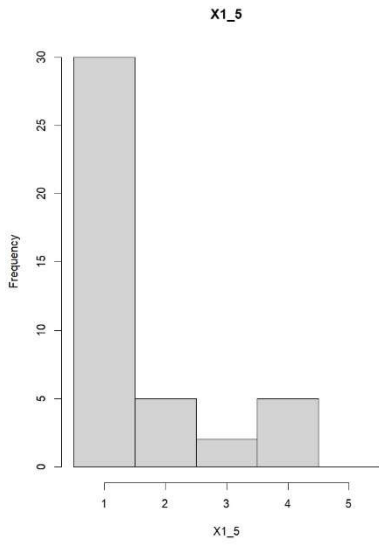
番号	状況	会話テキスト
1_1	状況：あなたとジョンは散歩しています	あなた：メルセデスを運転したことある？ ジョン：私は[高級車を絶対に運転]しないの
1_2	状況：あなたとジョンは散歩しています	あなた：メルセデスを運転したことある？ ジョン：私は[車を運転]しないの
1_3	状況：あなたとジョンは散歩しています	あなた：メルセデスを運転したことある？ ジョン：私は[贅沢]しないの
1_4	状況：あなたとジョンは散歩しています	あなた：メルセデスを運転したことある？ ジョン：私は[ヨーロッパには行か]ないの
1_5	状況：あなたとジョンは散歩しています	あなた：メルセデスを運転したことある？ ジョン：私は[散歩を絶対に]しないの
2_1	状況：あなたはカスミとクジ引きを作っています	あなた：ランダム化しておいた カスミ：[不正防止]に強い設計だね
2_2	状況：あなたはカスミとクジ引きを作っています	あなた：ランダム化しておいた カスミ：[統計的]に強い設計だね
2_3	状況：あなたはカスミとクジ引きを作っています	あなた：ランダム化しておいた カスミ：[物理的]に強い設計だね
2_4	状況：あなたはカスミとクジ引きを作っています	あなた：ランダム化しておいた カスミ：[運命]に強い設計だね
2_5	状況：あなたはカスミとクジ引きを作っています	あなた：ランダム化しておいた カスミ：[万有引力]に強い設計だね
3_1	状況：あなたはハルトと勉強しています	あなた：ついでにこの課題も ハルト：机の上に載る[課題]の数は数えたことある？
3_2	状況：あなたはハルトと勉強しています	あなた：ついでにこの課題も ハルト：机の上に載る[消しゴム]の数は数えたことある？
3_3	状況：あなたはハルトと勉強しています	あなた：ついでにこの課題も ハルト：机の上に載る[ホコリ]の数は数えたことある？
3_4	状況：あなたはハルトと勉強しています	あなた：ついでにこの課題も ハルト：机の上に載る[後悔]の数は数えたことある？
3_5	状況：あなたはハルトと勉強しています	あなた：ついでにこの課題も ハルト：机の上に載る[麒麟]の数は数えたことある？

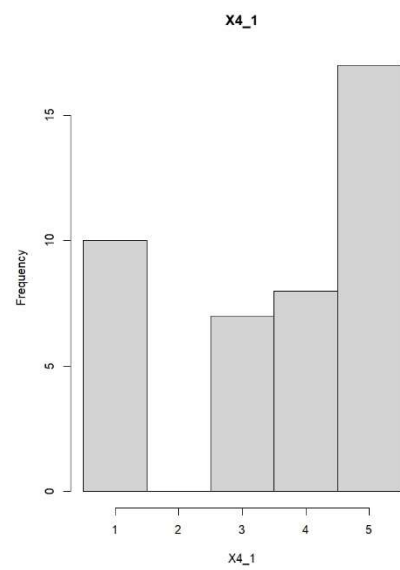
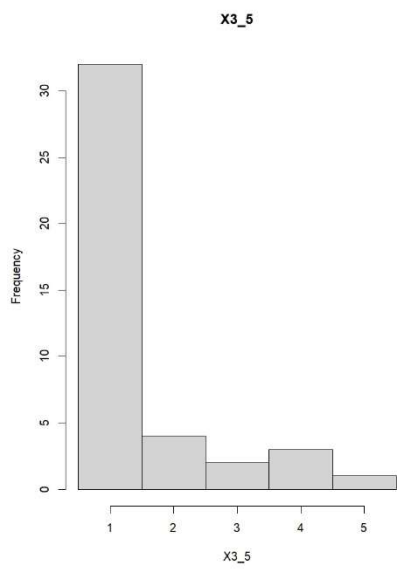
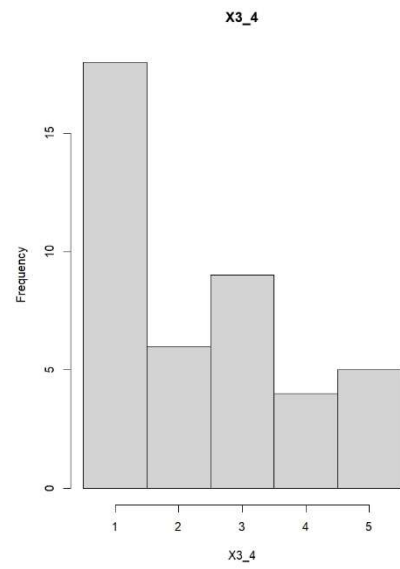
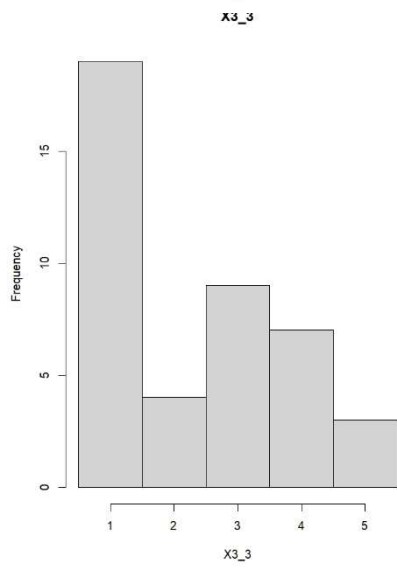
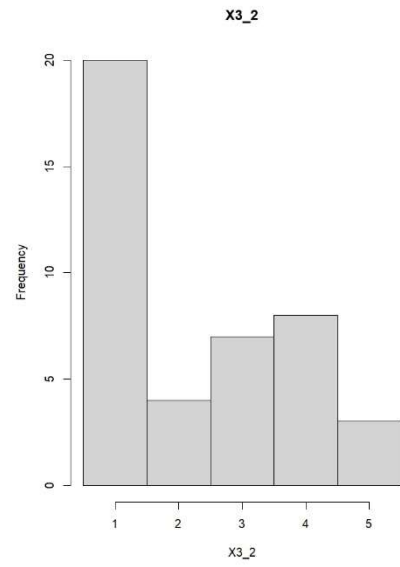
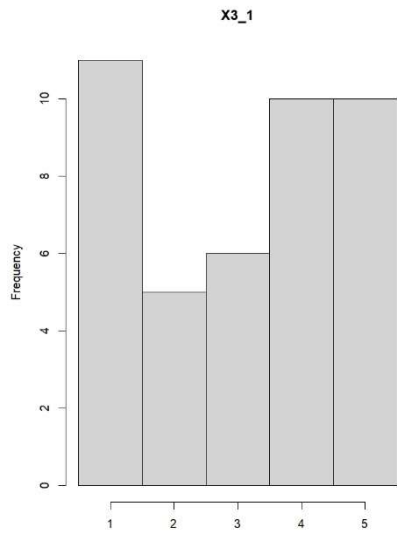
4_1	状況:あなたはサクヤとパーティーの打ち合わせを行っています	あなた: 最初からすべて話すね サクヤ: [地図]が一枚あると、道に迷わない気がする
4_2	状況:あなたはサクヤとパーティーの打ち合わせを行っています	あなた: 最初からすべて話すね サクヤ: [お守り]が一枚あると、道に迷わない気がする
4_3	状況:あなたはサクヤとパーティーの打ち合わせを行っています	あなた: 最初からすべて話すね サクヤ: [丈夫な靴]が一枚あると、道に迷わない気がする
4_4	状況:あなたはサクヤとパーティーの打ち合わせを行っています	あなた: 最初からすべて話すね サクヤ: [目隠し]が一枚あると、道に迷わない気がする
4_5	状況:あなたはサクヤとパーティーの打ち合わせを行っています	あなた: 最初からすべて話すね サクヤ: [カブトムシ]が一枚あると、道に迷わない気がする
5_1	状況:あなたとレイカは友達の誕生日プレゼントを選んでいきます	あなた: この花束はどうか？ レイカ: この花束は[季節の花]が使われているね
5_2	状況:あなたとレイカは友達の誕生日プレゼントを選んでいきます	あなた: この花束はどうか？ レイカ: この花束は[ラッピング]が使われているね
5_3	状況:あなたとレイカは友達の誕生日プレゼントを選んでいきます	あなた: この花束はどうか？ レイカ: この花束は[ミニマリズム]が使われているね
5_4	状況:あなたとレイカは友達の誕生日プレゼントを選んでいきます	あなた: この花束はどうか？ レイカ: この花束は[湿度]が使われているね
5_5	状況:あなたとレイカは友達の誕生日プレゼントを選んでいきます	あなた: この花束はどうか？ レイカ: この花束は[言語そのもの]が使われているね
6_1	状況:あなたとサトシはレストランでメニューを見えています	あなた: この揚げ物頼んでみようか サトシ: この揚げ物の[人気]は高いですね
6_2	状況:あなたとサトシはレストランでメニューを見えています	あなた: この揚げ物頼んでみようか サトシ: この揚げ物の[カロリー]は高いですね
6_3	状況:あなたとサトシはレストランでメニューを見えています	あなた: この揚げ物頼んでみようか サトシ: この揚げ物の[雰囲気ポイント]は高いですね
6_4	状況:あなたとサトシはレストランでメニューを見えています	あなた: この揚げ物頼んでみようか サトシ: この揚げ物の[偏差値]は高いですね
6_5	状況:あなたとサトシはレストランでメニューを見えています	あなた: この揚げ物頼んでみようか サトシ: この揚げ物の[物語性]は高いですね

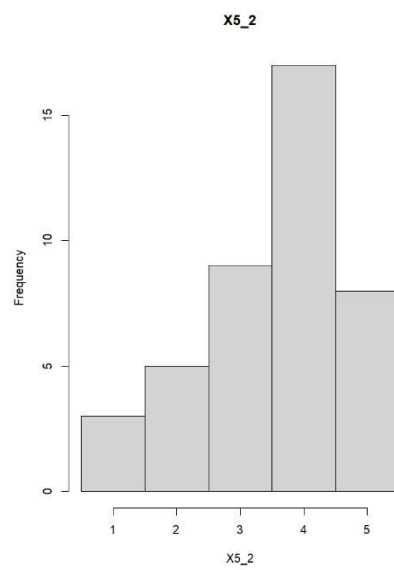
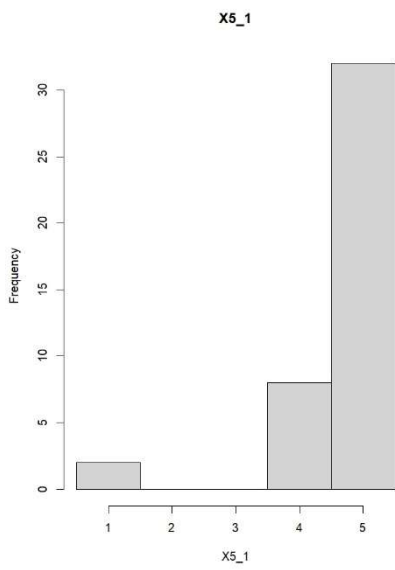
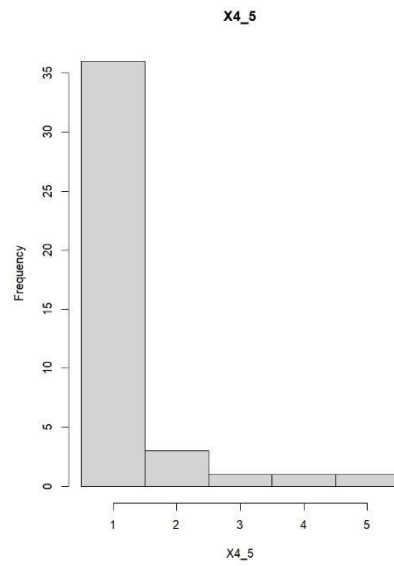
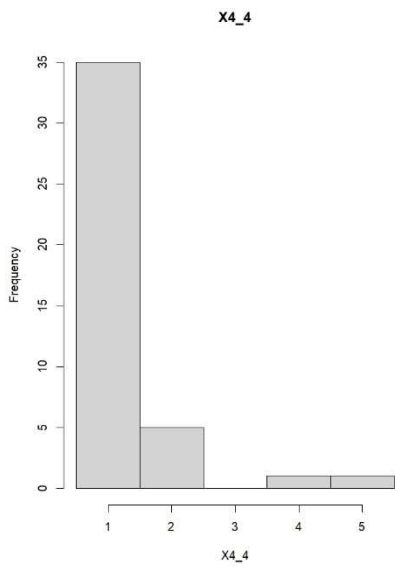
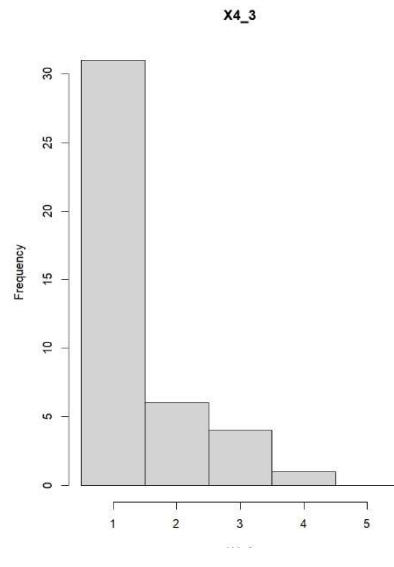
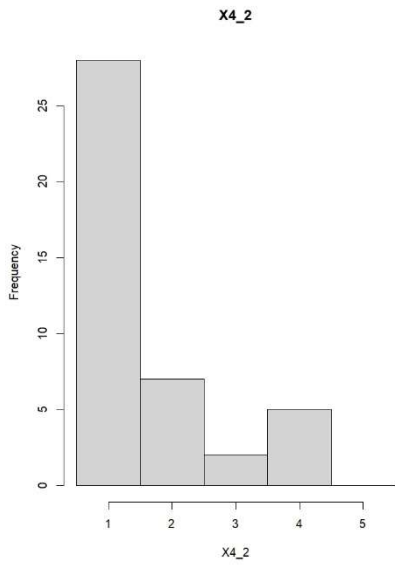
## 付録 B：理解容易度調査（3章）のために作成した会話テキストの理解容易度評定の分布

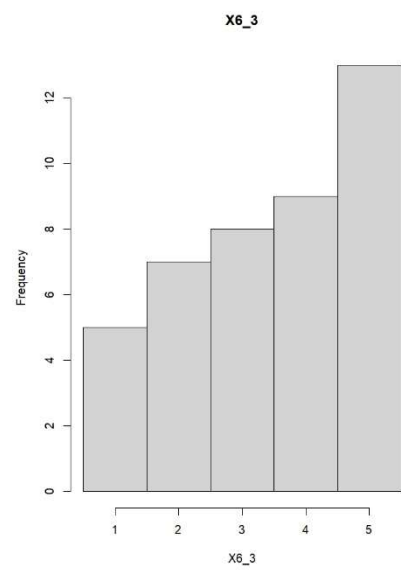
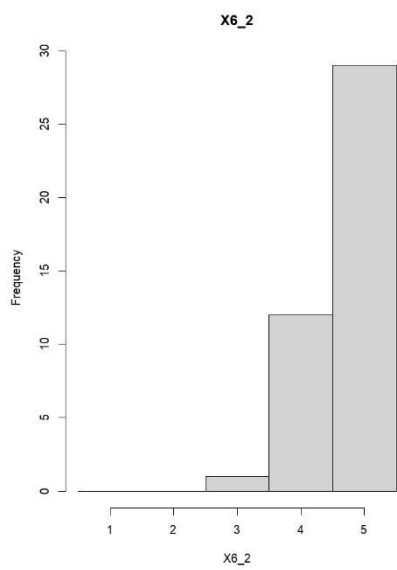
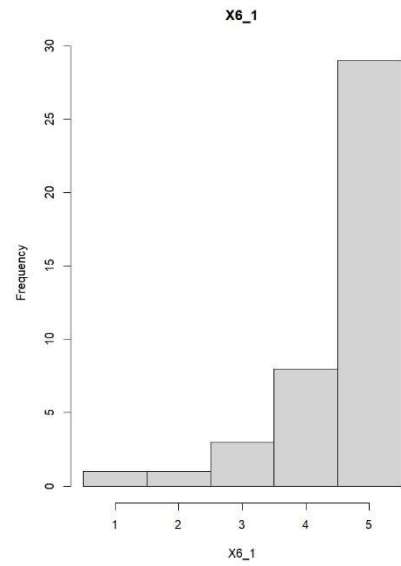
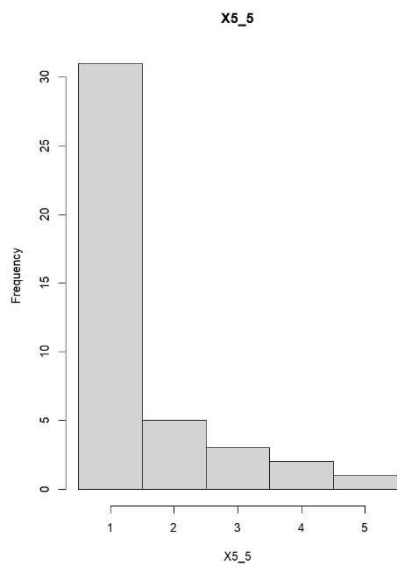
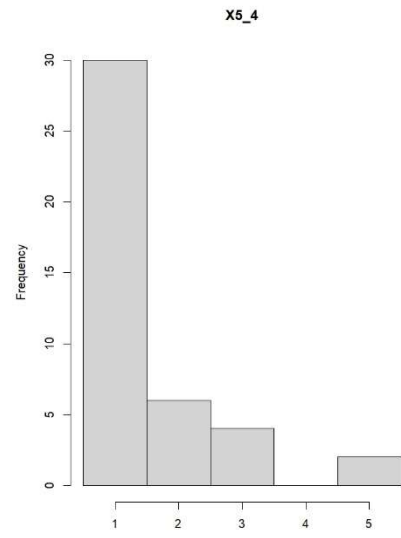
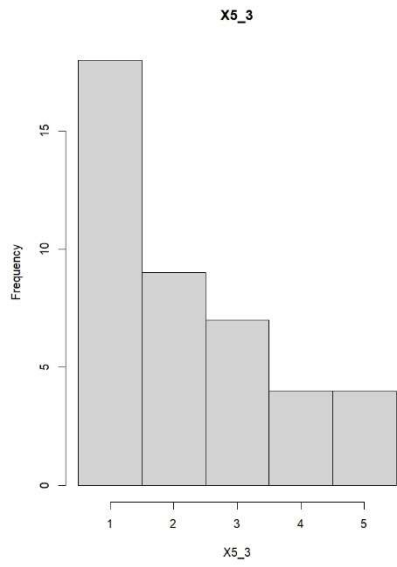
縦軸は回答数，横軸は理解容易度を示す．またタイトルは a\_b となっており，a が状況，b が理解容易評定値を示す．

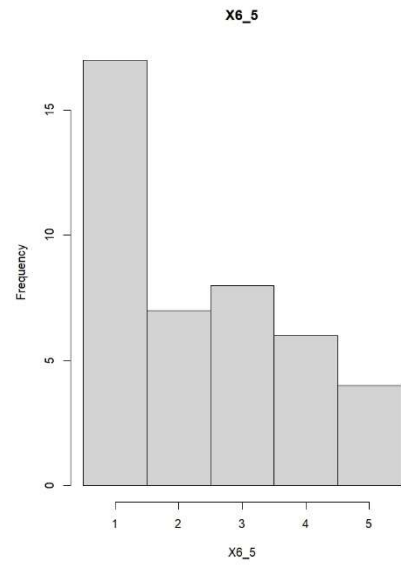
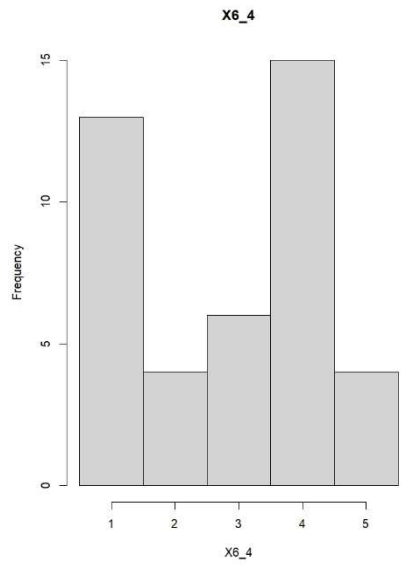












付録 C：各参加者に提示した刺激の順序

ID	順序 1	順序 2	順序 3	順序 4	順序 5	順序 6
101	A	F	C	B	D	E
102	D	A	F	E	B	C
103	B	A	F	D	E	C
104	B	D	C	A	E	F
105	B	F	E	C	A	D
106	D	B	E	A	C	F
107	D	B	E	F	C	A
108	A	E	F	B	D	C
109	B	E	D	A	F	C
110	F	B	C	A	D	E
111	E	D	F	A	B	C
112	F	E	C	D	B	A
113	E	C	D	F	A	B
114	F	A	D	B	E	C
115	E	F	B	D	A	C
116	C	E	A	F	B	D
117	E	F	D	A	C	B
118	B	D	F	A	C	E
119	D	F	E	C	A	B
120	C	F	D	E	A	B
121	F	D	A	B	C	E
122	F	E	B	C	D	A
123	E	C	F	B	D	A
124	D	F	B	E	A	C
125	A	D	F	C	B	E
126	F	B	D	C	E	A
127	D	E	F	A	C	B
128	B	A	D	C	E	F
129	F	E	A	B	C	D
130	C	F	B	D	A	E
131	D	E	C	F	B	A
132	D	A	E	C	F	B

133	E	A	C	B	F	D
134	B	C	D	F	E	A
135	E	C	A	D	B	F
136	A	D	B	E	C	F
137	C	B	A	D	F	E
138	F	D	E	C	A	B
139	C	B	A	D	E	F
140	B	C	A	E	F	D
141	C	D	A	F	E	B
142	A	C	E	B	D	F
143	A	B	C	E	F	D
144	A	C	B	F	D	E
145	C	A	B	E	F	D

---

## 付録 D：直感回答課題の提示スライド（刺激 A）

<p>設問1</p>	<p>[設問1] あなたから見て、相手は何を伝えたいのか すぐに一言で教えてください</p>
<p>状況：あなたとジョンは 散歩しています</p>	<p>状況：あなたとジョンは 散歩しています  あなた：メルセデスを運転したことある？</p>
<p>ジョンの返答を考えてください</p>	<p>状況：あなたとジョンは 散歩しています  あなた：メルセデスを運転したことある？ ジョン：私は高級車を絶対に運転しないよ</p>
<p>制限時間・15秒</p> <p>状況：あなたとジョンは 散歩しています  あなた：メルセデスを運転したことある？ ジョン：私は高級車を絶対に運転しないよ  ジョンの発言の意図は何ですか？ すぐにひと言で回答してください</p>	<p>回答をやめてください</p>

刺激 B～F も同様に提示

## 付録 E：多数回答課題の提示スライド（刺激 A）

<p>設問2</p>	<p>[設問2] あなたから見て、相手は何を伝えたいのか できるだけたくさん教えてください</p>
<p>状況：あなたとジョンは 散歩しています</p>	<p>状況：あなたとジョンは 散歩しています  あなた：メルセデスを運転したことある？</p>
<p>ジョンの返答を考えてください</p>	<p>状況：あなたとジョンは 散歩しています  あなた：メルセデスを運転したことある？  ジョン：私は高級車を絶対に運転しないよ</p>
<p>制限時間：3分</p> <p>状況：あなたとジョンは 散歩しています  あなた：メルセデスを運転したことある？  ジョン：私は高級車を絶対に運転しないよ  ジョンの発言の意図は何ですか？ できるだけたくさん教えてください 考えていることを1つ目は～。2つ目は～。という形で話し続けてください</p>	<p>回答をやめてください</p>

刺激 B～F も同様に提示

## 付録 F：解釈選択課題の提示スライド（刺激 A）

<p>設問3</p>	<p>【設問3】 あなたが考えた相手の意図を 一つずつ声に出して比べ 意図の解釈として一番適切だと感じるもの を選んでください</p>
<p>状況：あなたとジョンは 散歩しています</p>	<p>状況：あなたとジョンは 散歩しています あなた：メルセデスを運転したことある？</p>
<p>ジョンの返答を考えてください</p>	<p>状況：あなたとジョンは 散歩しています あなた：メルセデスを運転したことある？ ジョン：私は高級車を絶対に運転しないよ</p>
<p>状況：あなたとジョンは 散歩しています あなた：メルセデスを運転したことある？ ジョン：私は高級車を絶対に運転しないよ 右のモニターの画面に注目し回答してください</p>	<p>制限時間：3分 状況：あなたとジョンは 散歩しています あなた：メルセデスを運転したことある？ ジョン：私は高級車を絶対に運転しないよ ジョンの意図の解釈を 一つずつ声に出して比べ 一番適切だと感じるものを選んでください 『1つ目は〇だから××』 『2つ目は〇だと××』 『以上から、〇つ目の～である』 『理由は～だからだ』</p>
<p>回答をやめてください</p>	

刺激 B～F も同様に提示

## 付録 G：全参加者の Q1・3 の一致率

全参加者の直感回答課題（Q1）と解釈選択課題（Q3）の一致率

参加者 ID	比較可能 数	Q1・3 不一致 数	一致 率	推定一致 数	規則一致 数	認知的共感一 致数
101	6	1	0.83	3	2	0
104	6	2	0.67	2	2	0
105	5	1	0.80	3	1	0
106	6	0	1.00	5	1	0
109	4	1	0.75	2	1	0
110	6	2	0.67	2	2	0
111	6	0	1.00	4	2	0
112	5	2	0.60	2	1	0
113	5	0	1.00	4	1	0
114	6	2	0.67	2	1	1
115	3	0	1.00	0	3	0
116	6	1	0.83	3	2	0
117	5	1	0.80	3	1	0
118	6	2	0.67	1	2	1
119	6	2	0.67	1	3	0
120	4	0	1.00	3	1	0
121	5	2	0.60	2	1	0
122	5	3	0.40	2	0	0
123	6	1	0.83	2	2	1
124	6	1	0.83	3	2	0
125	5	3	0.40	2	0	0
126	5	2	0.60	3	0	0
127	4	0	1.00	4	0	0
128	6	1	0.83	2	3	0
129	6	2	0.67	3	1	0
130	2	0	1.00	1	1	0
131	6	3	0.50	1	1	1
132	6	3	0.50	3	0	0
133	6	1	0.83	5	0	0
134	6	0	1.00	4	2	0

135	6	0	1.00	4	2	0
136	4	2	0.50	2	0	0
137	6	3	0.50	0	3	0
138	5	0	1.00	3	2	0
139	4	0	1.00	4	0	0
140	5	0	1.00	4	1	0
142	6	1	0.83	2	3	0
143	6	4	0.33	1	0	1

---

## 付録 H：本実験で提示した書類

以下の順で提示する

- ・実験説明書
- ・同意書
- ・同意撤回書

なお、個人情報の部分は黒塗りである。

## 発話解釈に関する実験についての説明書

以下の事項をよく読み、同意いただける場合は別添の同意書に氏名を自署してください。

### 1. 実験参加者の権利

実験参加者となることは自由意思によるものであり、実験の実施前・実施中を問わずいつでも拒否することができ、また、実施後は 30 日以内であれば参加同意を撤回することができます。拒否・撤回することにより何ら不利益が生じることはありません。

### 2. 研究の目的と意義

本研究は、コミュニケーションにおける発話を人々がどのようにして解釈しているかについて、推論形式に着目して明らかにすることを目的とします。本研究を通じ、聞き手が文脈や状況に応じてどのように推論方略を選択しているのかについて、基礎的かつ実証的な知見を提供し、コミュニケーション不全の要因や対処方法を明らかにすることに繋がるという意義があります。

### 3. 研究のために拘束されるおおよその時間

本実験は、実験実施前・実施中の説明、回答練習、発話解釈課題、および、事後アンケートで構成され、最長で 2 時間 15 分程度かかります。拘束される時間に応じて、謝礼を本学の規定に従ってお支払いします(1 時間あたり 1300 円)。実験中には適宜休憩をはさみます。また、参加者からの申し出によりいつでも休憩を取ることができます。

### 4. 実験参加者が行うこと

参加者には2人の会話テキストを読み設問に回答していただきます。回答は、考えていることをすべて口に出して発話する方法(以下、「発話思考法」)で行っていただきます。その際、実験中に提示する画面と発話の音声をZoomにより記録させていただきます。なお、カメラによる写真・動画の撮影は一切いたしません。その後、事後アンケートに回答していただきます。

実験をスムーズに行うため、実験室入室後には、スマートフォンやスマートウォッチなどの電子機器の電源をお切りいただき、すべての荷物を実験室内に設置した籠に入れてくださるよう、ご協力をお願いいたします。

実験は次の3つのステップから構成されます。

- (1) 課題の説明と練習: 実験課題に慣れていただくため、説明動画を閲覧し、課題と発話思考法による回答の練習をしていただきます。所要時間は30分程度です。
- (2) 発話解釈課題: 会話テキストを読んで設問に発話思考法で回答していただきます。設問は、発言意図について直感的な解釈、様々な解釈、その中から最適と思う解釈の選択の3つです。会話テキストは全部で6つあります。所要時間は会話テキスト間の休憩も含めて1時間程度です。
- (3) 事後アンケート: 参加者ご自身の性格や普段のコミュニケーション、および、課題の評価についてのアンケートに回答していただきます。所要時間は15分程度です。

## 5. 予測される実験参加者への影響（不利益, 不快感等）

- ① 会話テキストは一般的に不快感が生じるようなものではありません。しかし、苦痛や不快を感じた場合はいつでも実験を取りやめることができます。
- ② 実験実施には、休憩を含めて 2 時間 15 分程度かかります。休憩は1つの状況・会話テキストごとに 2 分、すべての発話解釈課題終了後に 5 分ほど取ります。この他にも、申し出によりいつでも休憩を取ることができます。課題に取り組むことが苦痛や不快を与えるものであると判断される場合は、実験を取りやめることができます。
- ③ 実験に用いる機器の操作のため、実験実施中は参加者の近くに研究実施者がいます。これが心理的ストレスを感じるなどで苦痛や不快を与えるものであると判断される場合は、実験を取りやめることができます。
- ④ 参加者の回答について「良い・悪い」の評価は一概に行いません。また、個人の政治信条・その他信仰等と関連付けることはありません。

## 6. 予測される実験参加者及びそれ以外の人への利益

発話解釈のメカニズムを科学的に検討した成果を、学会や学術雑誌へ発表することを通して人類共有の知識として提供できることが期待されます。

## 7. 実験参加者を特定できる情報がどの程度守られるか

本実験による取得するデータは、Zoom で記録した課題提示画面の録画と回答音声の録音、Zoom の自動文字起こし機能により文字化した回答音声のテキスト、および、Google Form で取得したアンケート回答です。これらすべてのデータは、実験終了後 1 週間以内にセキュリティを保ったクラウドストレージ(J-Strage Box)内に保存し、それぞれ Zoom のクラウドストレージ、Google Drive 内から削除します。テキスト内に個人を特定できる情報が含まれる場合は、新たに符号又は番号を付けて匿名化します。録音データは技術的に匿名化が困難であるため、原本のまま J-Strage Box 内に保存しますが、流出・情報漏洩が生じないよう厳格に管理いたします。

データの開示・同意撤回に応じるために、データと参加者を対応づける対応表を作成します。対応表は実験データとは分けて SSD に暗号化して保存し、SSD は研究責任者の居室の鍵のかかる棚にて保管します。すべてのデータは研究期間終了時(2026 年 3 月末)から 10 年後に復元不可能な方法で廃棄します。

すべての実験データの処理はセキュリティを保ったコンピュータ環境の中で行われます。データの評価は科学的な検討のみに用いられ、いかなる個人情報も他へ漏洩することはありません。取得したデータ等は本人の同意なく研究責任者・研究実施者以外に渡ることはありません。本研究を通じて得られた情報は、貴重な研究成果として、国内学会・国際会議・論文誌・修士論文等、およびアウトリーチ活動で発表されることがあります。いずれの場合も、参加者の名前や個人が特定される情報は一切示さず、プライバシーの保護に十分配慮いたします。

## 8. データの開示

実験参加者からのデータ開示の請求には、研究期間終了日である 2026 年 3 月 31 日まで応じます。研究期間終了後は対応表を廃棄するため、連結不可能匿名化となりますので、実験参加者のデータは特定できなくなり開示の請求に対応できません。

9. 利益相反, 倫理相反の配慮

北陸先端科学技術大学院大学「人を対象とした研究の実施に関する規則」に従い, 参加者より書面上での同意を得たうえで実験を実施します. 実験参加の有無・同意の撤回や実験での行動が, 学業評価および研究指導に影響することは一切ありません. また, いかなる企業・団体等との間に利益相反に相当する事項はありません.

10. 同意撤回について

本実験の参加への同意は, 実験実施中はいつでも, また, 実験実施後は 30 日以内まで, 文書により撤回することができます. 同意撤回書を研究実施者に直接あるいはメールにより提出することにより, 参加同意は撤回され, 提供いただいたデータを廃棄いたします. ただし, 実験終了後 31日以降は学会発表・論文発表を予定しているため, データの廃棄は不可能となることにご留意下さい. また, 同意を撤回しても, そのことにより参加者に何ら不利益はありません.

11. 実験内容の守秘について

本研究の目的・意義を達成できるように, 実験の内容を他者に話したり実験説明書を開示したりしないでください. もし開示が必要な場合は研究期間終了日(2026 年 3 月 31 日)以降にお願いいたします.

12. 研究内容, 倫理・安全面等についての質問に答えることができる担当者の連絡先

<p>■研究内容についての問い合わせ先:</p> <p>研究実施者:</p> <p>〒923-1211石川県能美市旭台1-1 北陸先端科学技術大学院大学 先端科学技術研究科 橋本研究室博士前期課程2年 松崎 由幸 [Redacted]</p> <p>研究責任者:</p> <p>北陸先端科学技術大学院大学 先端科学技術研究科 教授 橋本 敬 [Redacted] [Redacted]</p>	<p>■人権ハラスメントに関する問い合わせ先:</p> <p>先端科学技術研究科ハラスメント相談員:</p> <p>中分 遥 准教授 [Redacted]</p> <p>キム ウニョン(Kim Eunyoung) 准教授 [Redacted]</p>
	<p>■研究倫理に関する先端科学技術研究科相談員:</p> <p>池田 満 教授 [Redacted]</p>

北陸先端科学技術大学院大学  
先端科学技術研究科  
教授 橋本 敬

## 同意書

北陸先端科学技術大学院大学  
所属：先端科学技術研究科  
研究責任者名：橋本 敬 殿

私は、「発話解釈に関する実験」の研究・実験の実施について、研究実施者 松崎 由幸 より、  
年 月 日、説明書を用いて知識科学系研究棟Ⅱ棟8階（K-83）で説明を受け、研究計画の目的、意義、方法、個人情報保護の方法、安全管理での配慮などについて十分理解しましたので、計画に参加し、求められた私個人に係る情報、データ等を提供することに同意いたします。

説明を受けて理解した項目

(理解できた項目の左に○、理解できなかった項目の左に×を記載ください。)

- |   |
|---|
| <p>1 研究計画の概要に関する事項<br/>( )・研究の目的、意義<br/>( )・提供する情報、データ等</p> <p>2 個人情報保護の方法に関する事項<br/>( )・個人情報の収集が、研究目的、研究計画に照らして必要であること<br/>( )・提供を受けたデータ等処理の匿名化の方法<br/>( )・データの保管・管理について適切になされること</p> <p>3 侵襲および安全管理に関する事項<br/>( )・予想される不快感、負担等（実験参加のために拘束されることによる不快感）</p> <p>4 インフォームド・コンセントに関する事項<br/>( )・研究計画への参加は任意であること<br/>( )・研究計画に参加に同意しないことにより不利益な対応を受けないこと<br/>( )・研究計画に参加に同意した後でも、実験終了後 30 日以内であればいつでも文書により同意を撤回することができるが、実験終了後 31 日以降はデータの廃棄は困難であること<br/>( )・同意を撤回しても、そのことにより何ら不利益を蒙らないこと<br/>( )・同意を撤回した場合、提供されたデータ等は廃棄されること<br/>( )・収集したデータ等は、本人の同意を得ることなく他者に渡さないこと<br/>( )・研究成果の発表の方法について、学会発表、論文発表の予定<br/>( )・研究計画参加に対して謝礼を支払うこと</p> |
|---|

年 月 日

氏 名 (自署)

---

# 同意撤回書

研究責任者 橋本 敬 殿

私は、\_\_\_\_年\_\_\_\_月\_\_\_\_日付で同意した「発話解釈に関する実験」への参加の同意を撤回いたします。

本人の署名欄

本人署名

\_\_\_\_\_

署名日

\_\_\_\_\_