

雑談の応答生成に向けた修辞関係の分析

徳久良子 下岡和也 寺嶋立太

(株)豊田中央研究所

{tokuhisa,shitaoka,ryuta}@mosk.tytlabs.co.jp

1 はじめに

従来の課題遂行対話システムでは、タスク達成に必要な項目を記述したフォームを利用して大局的に対話を制御する手法が用いられてきた [3, 8, 10]。例えば、Goddeauらは「車メーカー、車体の色、値段」などの項目を持つフォームを利用して、ユーザの好みの車を提示する対話システムを実現している。一方、非課題遂行対話（便宜的に雑談と表現する）システムでは、明示的なタスクが存在しないため、ユーザから入手すべき項目をあらかじめ定義して大局的に対話を制御することは難しい。そこで我々は、大局的な構造を仮定せず局所的な発話のやりとりにより雑談を実現することを目指す。

少数の研究事例ではあるが、局所的な発話の関係に基づき応答を生成する手法が提案されている [4, 5]。Greenらは、質問と応答との修辞関係を利用して付加情報を応答する質問応答システムを実現した。具体的には、質問と応答との修辞関係の分析に基づき応答生成ルールを定義し、それによって応答を生成することで、ある事象についての質問に対して、修辞関係を持つ事象を応答するシステムを構築している。以下に、Greenらが実現した質問応答システムの対話例を示す。

User: Is Mark at the office?

System: a) No.

b) He's at home. *<Contrast>*

上記 b では、質問の内容と修辞関係 (*Contrast*) を持つ事象によって応答することで、質問に対して付加情報を提示している。

さらに近年では、大量のテキストを利用して、修辞関係を持つ事象を自動獲得する研究も行なわれている [2, 7, 6]。これを応答生成に応用すれば、修辞関係を持つ事象を自動獲得した上で、事象間の修辞関係を利用して応答を生成することができる。しかし、どのような修辞関係が雑談の応答生成に有効かが自明でない上に、あらゆる修辞関係 (*Contrast, Causal relation* など) について、修辞関係を持つ事象を精度良く自動獲得できるとは言えない。

我々はこれまでに、人間同士の課題遂行対話と雑談とを比較し、雑談に特徴的なやりとりについて分析した。まず、課題遂行対話と雑談における発話の繋がり（連続長（発話連鎖長と呼ぶ））を調べたところ、雑談では課題遂行対話より発話連鎖長が1%水準で有意に長いことが分かった ($t=5.41, p<.01$; 表 1 参照)¹。さらに、発話連鎖を起す発話について分析したところ、間接応答と問い返しが多く出現することが分かった。間接応答とは他者の質問に対して付加情報を提示する応答で、問い返しとは他者の働きかけ発話に関して質問する応答を指す（詳細は文献 [11] 参照）。

表 1: 課題遂行対話と雑談における発話連鎖長

	発話連鎖長							
	1	2	3	4	5	6	7	8
課題遂行対話	38	27	9	0	0	0	0	0
雑談	119	124	47	26	12	5	4	2

本稿では、雑談に特徴的に出現する間接応答と問い返し、どのような修辞関係によって構成されるかを明らかにする。以下、2章では修辞関係の注釈づけの方法を説明し、3章で雑談に特徴的な修辞関係を分析する。最後に、4章でまとめと今後の課題を述べる。

2 修辞関係の注釈づけ

2.1 雑談コーパスの収集

雑談における修辞関係の分析のため、我々は、面識のない2名の対面対話により雑談を収集した [12]。収録は、話者 A (20代と60代男女、計4名) と話者 B (30代女性) が対面で行った。雑談収録に先立ち、話者 A は15枚の話題カード (e.g. 食べ物, 旅行, スポーツ, 趣味) から3枚のカードを選択した。ただし、話題カードは、雑談開始のトリガとして用い、対話進行中に話題がカードの内容からそれともよいものとした。このような方法で、15~20分間の雑談が収録された時点で対話を終了した。その結果、図 1 および図 3 に示すような雑談を収録した。

¹我々の分析では、[働きかけ-応答]の発話連鎖長は1、[働きかけ-応答-了解]の発話連鎖長は2とした。

2.2 修辞関係の注釈づけの方法

本節では、先行発話と間接応答および先行発話と問い返しとの修辞関係の注釈づけの方法を述べる。

2.2.1 間接応答に対する注釈づけ

Green らは、間接応答における修辞関係を、Rhetorical Structure Theory (RST) の Rhetorical Relation (RR) を簡略化した修辞関係に基づいて分析した [4]。我々は、できるだけ詳細に質問と間接応答との修辞関係を分析するため、RST で定義された 23 種類の RR を全て用いて修辞関係の分析を行なった。表 2 に RST で定義された RR の一覧を示す。

表 2: RST における Rhetorical Relation の一覧

Circumstance	Antithesis	Concession
Solutionhood	Elaboration	Background
Enablement	Motivation	Evidence
Justify	Volitional Cause	Non-Volitional Cause
Purpose	Volitional Result	Non-Volitional Result
Condition	Otherwise	Interpretation
Evaluation	Restatement	Summary
Sequence	Contrast	

RST の RR とは本来、節間の修辞関係を表現する体系である。ここでは質問文と間接応答との間に RR を付与するため、以下の手順でタグ付与作業を行なう。

手順 1) 質問文を平叙文に言い換える。

手順 2) 手順 1 で作成した平叙文と間接応答間の修辞関係を表 2 に示す RR の定義からひとつ選択し付与する。ただし、平叙文と間接応答が複数の節から成る場合は、主節間の修辞関係のみを付与する。

図 1 に、質問と間接応答との間に RR を付与した例を示す。図 1 中の DR は直接応答を、IR は間接応答を示す。図 1 の例 1 に RR を付与する場合、まず、手順 1 では「お仕事に行ってるんですか?」という質問を「お仕事に行っちゃいます」という平叙文に言い換える。次に、手順 2 で、S1 (お仕事に行っちゃいます) と S2 (それで保育園に預けてあるのよ) との修辞関係を考える。この場合は、S2 は S1 の意志性のある結果なので Volitional Result を付与する。

2.2.2 問い返しに対する注釈づけ

前述した通り、RST の RR とは本来、節間の修辞関係を表現する体系である。従って、例えば以下のような先行発話と問い返しの関係は RST の RR の定義では表現できない。

話者 A: 駅前に新しい店ができたよ。

話者 B: そうなの?

そこで我々は、修辞関係を分析するための前処理として、Rieser らの定義した relation antecedent に従って先行発話と問い返しとの関係を分析した [9]。Rieser らの定義した relation antecedent とは、先行発話と問い

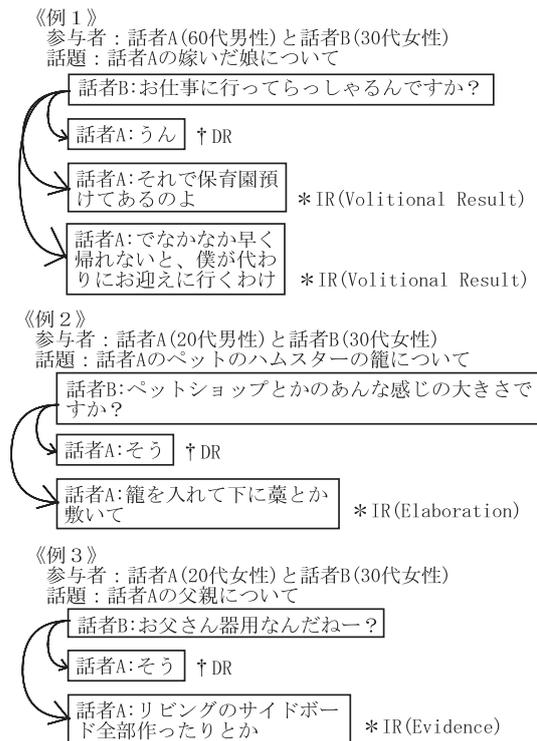


図 1: 質問と間接応答との修辞関係の分析例

返しとの表層的な関係を分析するもので、表 3 に示す分類から成る。雑談コーパスに出現した 144 事例の問い返しを Rieser らの方法で分析した結果を図 2 に示す。

次に、図 2 で Addition が付与された 93 事例について、先行発話と問い返しとの間の修辞関係を分析する。修辞関係の分析には、RST の RR に POC(Part Of Clause) というタグを追加した、24 種類のタグを用いた。以下に付与手順を示す。

手順 1) 問い返しを平叙文に言い換える。

手順 2) 手順 1 で作成した平叙文が、先行発話の節の一部である場合には POC を付与する。平叙文が先行発話の節の一部でない場合には手順 3 に進む。

手順 3) 先行発話と手順 1 で作成した平叙文との修辞関係を表 2 に示す RR からひとつ選択し付与する。ただし、先行発話および問い返しが複数の節から成る場合は、主節間の修辞関係のみを付与する。

図 3 に、先行発話と問い返しとの間に RR を付与した例を示す。図 3 中の CR は問い返しを示す。図 3 の例 1 では、まず、手順 1 で「二人で?」という問い返しを平叙文「二人で」に言い換える。次に、手順 2 で「二人で」が先行発話の節(「中学生になったばかりの娘を

表 3: 先行発話と問い返しとの表層的な関係の定義

分類	定義	対話例
Addition	先行発話に情報付加した問い返し	話者 1: ますます家族が映画にはまって 話者 2: だって手軽に家で見れるんだもんね?
Repetition	先行発話の繰り返しによる問い返し	話者 1: 幼稚園受験するので 話者 2: 幼稚園受験?
Reformulation	先行発話の言い換えによる問い返し	話者 1: 入園するだけで大変で 話者 2: 入るのが?
Repet-Add	先行発話を繰り返し、さらに情報を付加した問い返し	話者 1: キャンプ行くの 話者 2: 毎年キャンプ行くの?
Independent	「え?」「そうなの?」などの特有の表現による問い返し	話者 1: どちらかというとお母さんが一生懸命 話者 2: そうなの?

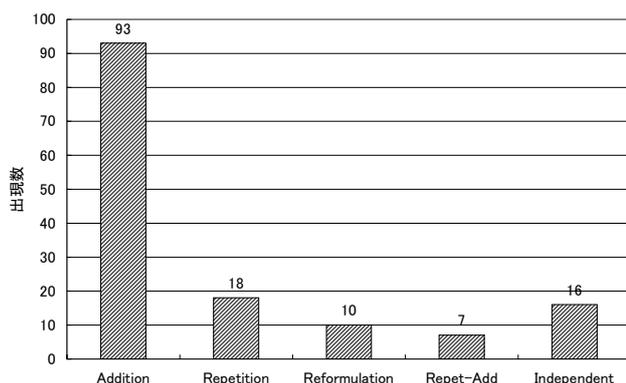


図 2: 先行発話と問い返しとの表層的な関係

置いて」もしくは「どっかいっちゃん」の一部かどうかを判断する。ここでは「どっかいっちゃん」と「二人で」から「二人でどっかいっちゃん」というひとつの節が構築できるため、POCを付与する。また、図3の例2では、手順1で「それはやっぱり勢い良く回しているからですか?」を「それはやっぱり勢い良く回しているからです」という平序文に言い換える。さらに、手順2で「それはやっぱり勢い良く回しているからです」が先行発話の節の一部かどうかを判断する。この場合は、いずれの節の一部でもないため手順3に進む。手順3では、先行発話の節S1(起きちゃうぐらいうるさい時もあるんですね)とS2(それはやっぱり勢い良く回しているからです)との修辞関係を考える。その結果、S2はS1の意志性のない原因であるため、Non-Volitional Causeを付与する。

2.3 修辞関係付与作業の信頼性評価

修辞関係タグ付与作業の再現可能性評価のため、節の雑談コーパスから間接応答と問い返しを各50事例抽出し、会話分析の経験のある作業員2名(作業員1, 作業員2と呼ぶ)が独立に修辞関係を付与した。

表4に、Cohenの κ を用いて評価した結果を示す[1]。一般に、 κ 値が0.6-0.8の場合は、タグ付与作業の一致率は比較的高いとされる。従って、表4から、間接応

《例1》
 参与者: 話者A(20代女性)と話者B(30代女性)
 話題: 話者Aの両親の仲の良さについて
 話者B: 中学生になったばかりの娘を置いてどっか
 いっちゃん
 話者A: 二人で? † CR(POC)

《例2》
 参与者: 話者A(20代男性)と話者B(30代女性)
 話題: 話者Aのペットのハムスターと運動器具について
 話者A: 可愛いんですけど起きちゃうぐらいうるさい
 時もあるんですね
 話者B: それはやっぱり勢い良く回しているからですか? † CR(Non-Volitional Cause)

《例3》
 参与者: 話者A(60代女性)と話者B(30代女性)
 話題: 話者Aのペットの犬と猫について
 話者A: 私は犬の世話はやらないから
 話者B: え、家の中のマルチーズも外のゴールデンの世話も全く
 やらないんですか? † CR(Elaboration)

図 3: 先行発話と問い返しとの修辞関係の分析例

答や問い返しに対する修辞関係は比較的安定して付与できると考えられる。

表 4: 修辞関係付与作業の信頼性評価結果

	事例数	κ 値
先行発話と間接応答との修辞関係	50	0.67
先行発話と問い返しとの修辞関係	50	0.76

3 間接応答と問い返しの修辞関係分析結果

2.3節の予備調査の結果から、修辞関係は比較的安定して付与できることが確認された。そこで、2.3節のタグづけ作業のうち1名により、先行発話と間接応答および問い返しの修辞関係を分析した。表5に質問と間接応答の、表6に先行発話と問い返しの修辞関係の分析結果を示す。

まず、質問と間接応答との修辞関係の出現頻度を表5に示す。因果関係に関する修辞関係(表5中†)をひとつの修辞関係と考えると、表5から、上位5個の修辞関係²により86.7%の間接応答を表現できることが分か

²Elaboration, Causal relation(表5中†), Evidence, Antithe-

る．具体的には，Elaboration の例として，図 1 の例 2 では質問文中の「あんな感じ」の詳細を間接応答で述べている．また，Evidence の例として，図 1 の例 3 では質問文中の「器用」な証拠を間接応答で述べている．

次に，先行発話と問い返しとの修辞関係を表 6 に示す．因果関係に関する修辞関係(表 6 中 †)をひとつの修辞関係と考えれば，表 6 から，上位 5 個の修辞関係³により 72.1%の問い返しを表現できることが分かる．具体的には，因果関係に関する修辞関係の例として，図 3 の例 2 では，先行発話の意志性のない原因を問い返している．また，Elaboration の例として，図 3 の例 3 では，先行発話中「犬の世話はやらない」の詳細を問い返している．

表 5: 質問と間接応答との修辞関係結果
† は，因果関係を表す Rhetorical Relation を示す．

Rhetorical Relation の分類	出現数	出現割合 (%)
Elaboration	27	27.6
Evidence	17	17.3
Volitional Result †	14	14.3
Antithesis	10	10.2
Evaluation	9	9.2
Circumstance	5	5.1
Volitional Cause †	5	5.1
Sequence	3	3.1
Background	2	2.0
Non Volitional Cause †	2	2.0
Enablement	1	1.0
Justify	1	1.0
Motivation	1	1.0
Non Volitional Result †	1	1.0
その他の Rhetorical Relation	0	0.0

表 6: 先行発話と問い返しとの修辞関係結果
† は，因果関係を表す Rhetorical Relation を示す．

Rhetorical Relation の分類 + POC	出現数	出現割合 (%)
Elaboration	16	17.2
POC	9	9.7
Contrast	9	9.7
Volitional Cause †	9	9.7
Interpretation	8	8.6
Volitional Result †	7	7.5
Antithesis	6	6.5
Evaluation	6	6.5
Evidence	5	5.4
Non Volitional Cause †	5	5.4
Circumstance	4	4.3
Non Volitional Result †	4	4.3
Sequence	3	3.2
Concession	1	1.0
Summary	1	1.0
その他の Rhetorical Relation	0	0.0

今回，質問と間接応答との修辞関係については全 23 種類，先行発話と問い返しとの修辞関係については全 24 種類の分類を用いて修辞関係の分析を行った．その

sis, Evaluation を指す．

³Causal relation(表 6 中 †), Elaboration, POC, Contrast, Interpretation を指す．

結果，間接応答や問い返しにおける修辞関係の出現には偏りが観察された．したがって，表 5 や表 6 で出現の多い修辞関係を中心に修辞関係を持つ事象を獲得することで，雑談の応答生成に必要な事象が効率的に整備できると考える．

4 まとめと今後の課題

本稿では，雑談に特徴的に出現する間接応答および問い返しに関して，先行発話との修辞関係を調べた．その結果，間接応答では Elaboration, Causal relation, Evidence などが，問い返しでは Causal relation, Elaboration などの関係が多く出現することが分かった．今回の分析で明らかとなったこれらの関係を中心に修辞関係を持つ事象を獲得することで，雑談システムの応答生成に応用可能な事象を効率的に整備できると考える．

今後は，今回の分析結果を踏まえて修辞関係を持つ事象を獲得すると共に，修辞関係を持つ事象を用いた雑談応答生成システムを構築する．

参考文献

- [1] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 1960.
- [2] Roxana Girju. Automatic Detection of Causal Relations for Question Answering. In *Proceedings of the ACL Workshop on Multilingual Summarization and Question Answering*, 2003.
- [3] David Goddeau, Helen Meng, Joe Polifroni, Stephanie Seneff, and Senis Busayapongchai. A FORM-BASED DIALOGUE MANAGER FOR SPOKEN LANGUAGE APPLICATIONS. In *Proceedings of the ICSLP*, 1996.
- [4] Nancy Green and Sandra Carberry. Interpreting and Generating Indirect Answers. *Computational Linguistics*, Vol. 25, No. 3, 1999.
- [5] Susan M. Haller. Planning for Intentions with Rhetorical Relations. In *Proceedings of the ACL Workshop on Intentionality and Structure in Discourse Relations*, 1993.
- [6] Kathleen Hornsby and Suzannah Hall. Generating Linear Orders of Text-Based Events. *HLT-NAACL 2004: Workshop on Computational Lexical Semantics*, 2004.
- [7] Christopher S.G. Khoo, Syin Chan, and Yun Niu. Extracting Causal Knowledge from a Medical Database Using Graphical Patterns. In *Proceedings of the ACL*, 2000.
- [8] Hiroaki Kitano and Carol Van Ess-Dykema. Toward a Plan-Based Understanding Model for Mixed-Initiative Dialogues. In *Proceedings of the ACL*, 1991.
- [9] Verena Rieser and Johanna Moore. Implications for Generating Clarification Requests in Task-Oriented Dialogues. In *Proceedings of the ACL*, 2005.
- [10] Amanda Stent, John Dowding, Jean Mark Gawron, Elizabeth Owen Bratt, and Robert Moore. The CommandTalk Spoken Dialogue System. In *Proceedings of the ACL*, 1999.
- [11] 徳久良子, 寺島立太. 雑談と課題遂行対話におけるやりとりの比較. 言語処理学会年次大会発表論文集, 2006.
- [12] 徳久良子, 寺島立太. 雑談における発話のやりとりと盛り上がりとの関連. 人工知能学会誌, Vol. 21, No. 2, 2006.