

話し言葉対話コーパスにおける協調的対話原則の分析と応用

堂坂 浩二 島津 明

NTT 基礎研究所

1 はじめに

話し言葉対話は、限られた推論能力しかもたない複数の自律的な主体が参加する実時間の活動である。それゆえ協調的な対話における話し言葉生成システムは困難な状況に直面することになる。システムは時間制限の下で発話を生成する必要がある、限られた推論能力しかもたない対話相手はシステムの発話を必ずしも理解してくれとは限らない。また、自律的な主体である対話相手はシステムの発話途中に応答や割り込みを行う。我々は、この困難な状況に対処することができる話し言葉生成システムとして、時間制限の下で漸次的に発話を生成し [2, 3], 対話相手の応答に臨機応変に対処するシステムを提案してきた [4].

こういった話し言葉生成システムが対話の中で適切に振舞うためには、システムが協調的な対話の中で従うべき原則が必要である。そのような原則を協調的対話原則と呼ぶ。本稿では、話し言葉対話コーパスにおいて発話単位、対話相手からの応答、応答に対処するための談話戦略の間の相互関係を統計的検定法を使って分析し、その分析結果に基づいて協調的対話原則を導く。続いて、協調的対話原則が話し言葉生成システムにおいて果たす重要な役割について論じる。協調的対話原則の効果は、対話シミュレーション実験によって確認されている。

2 関連研究

漸次的な言語生成のためのシステムやモデルが提案されているが、従来の多くのシステムやモデルは漸次的な文生成に注目しているか [1, 7], 書き言葉談話の漸次的な生成を目的としており [10], 話し言葉特有の談話構造は考慮されていない。話し言葉特有の談話構造に基づく漸次的生成モデルも提案されているが [2, 3], 対話相手からの応答には対処できない。

対話相手からの応答に対処することができる言語生成システムが提案されている [9]。このシステムは、ユーザがシステム発話を承認できないことを示す応答をした場合に、システム発話のどの部分が失敗したかを判断し、失敗した発話と同じ効果をもつ別の発話を生成する。しかし、このシステムは話し言葉対話の特徴を考慮していないため、話し言葉対話に適用するには不十分である。第 3 節に示すように、話し言葉対話では、相手からの応

答がない場合でもシステムが談話戦略を変更しなければならないことがある。

3 コーパスによる協調的対話原則の分析

3.1 発話単位

対話コーパスは次の対話実験により収集した。実験には 90 人の被験者が参加した。各対話において、二人の被験者 (N と E) には電話で会話することにより「 N がある場所から別の場所へ行く」という問題の解を見つけることを要請した。 E は問題を解くために十分な知識をもっており、 N はもっていないように選んだ。80 対話を収集し、20 対話を分析のために選んだ。

発話単位を次のように定義した: (1) 節は発話単位である, (2) 間投的表現 (例: 「はい」) は発話単位である, (3) つなぎ語と言ひ直しは発話単位の区切りである。20 対話には 2629 個の発話単位が現れた。発話単位の統語カテゴリは、節、名詞句 (または後置詞句) の 1 つ以上の並び、間投的表現、接続詞類、その他の機能語のうちいずれかであった [3]。節として実現された単位を節単位と呼び、名詞句または後置詞句の 1 つ以上の並びとして実現された単位を名詞単位と呼ぶ。1144 個の節単位と 436 個の名詞単位があり、それぞれの単位の中に現れる名詞句は 1 単位あたり平均 1.01 個であった。従属名詞句をもつ名詞句は全体で 1 個と数えた。この結果は、話し言葉では小さな発話単位が頻出することを示している。

情報の新旧の構造が発話単位の構成に影響を与えることはよく知られている [5]。発話単位の中の情報構造を調べるために、節単位と名詞単位の中に現れる名詞句を旧情報を担う名詞句と新情報を担う名詞句とに分類した。旧情報を担う名詞句として認定されたのは、代名詞、指示的な名詞句 (例: 「このバス」), 既に導入された対象を参照する提題化された名詞句 (例: 「愛甲石田は」), 直前に記述された命題や行為の中で言及された対象を参照する名詞句である。他の名詞句は新情報を担う名詞句とした。発話単位の中に含まれるすべての名詞句が旧情報を担うものなら、その単位を旧情報を担う単位と呼ぶ。そうでないなら、新情報を担う単位と呼ぶ。1580 個の名詞単位と節単位の中で、60 個 (4%) の発話単位だけが 2 個以上の新情報を担う名詞句を含んでいた。したがって次の原則が導かれる。

情報の新旧	名詞単位	節単位
旧情報	100 (73.1) 9.9	227 (253.9) 2.8
新情報	162 (188.9) 3.8	683 (656.1) 1.1
テーブル χ^2 値		17.7
棄却域 (確率 = 0.05)		$3.84 < \chi^2$

表 1: 発話単位の情報の新旧と統語カテゴリとの関係

原則 (p1) 新情報を担う名詞句を高々 1 つだけ含む発話単位を使え。

3.2 話し言葉対話における発話単位の役割

話し言葉コミュニケーションにおける節単位と名詞単位の役割を調べるために分析を進めた。以下の分析は、対話の統御 [6, 11] をもっていない対話参加者の発話は対象としない。また対話の統御をもつ参加者の発話であっても、相手の応答に対処するために発話されたものは対象としない。さらに、発話直後に修復された発話は対象としない。

まず、発話単位の文法カテゴリと情報の新旧という 2 つの属性の関係を χ^2 検定を使って分析した。なお、対話コーパスにおいて名詞句をともなっていることがない節単位 (例: 「分かりました」) は分析対象から外した。そのような単位は必然的に節単位のみで実現せざるを得ないので、名詞単位と節単位との間の選択とは関係がないと思われるからである。結果を表 1 に示す。テーブル χ^2 値が棄却域内にあるということは、2 つの属性が密接な関係にあることを示す。また、個々の属性値の組合せに関して、観測度数、期待度数、 χ^2 値が示されている。例えば、旧情報と名詞単位の組合せに関して、観測度数 100、期待度数 73.1、 χ^2 値 9.9 である。大きな χ^2 値をもつ属性値の組合せに関して、観測度数が期待度数よりも大きいならその組合せは生じやすく、観測度数が期待度数よりも小さいならその組合せは生じにくいことがわかる。したがって次の原則を導ける。

原則 (p2) 旧情報のみを伝達するために名詞単位を使え。

次に、発話単位の統語カテゴリと情報の新旧が発話単位への応答の有無に与える影響を調べた。応答としては、間投的表現、発話単位の復唱、発話単位を確認するための発話、質問に対して応答するための発話などがある。結果を表 2 に示す。まず、この結果からわかることは、新情報を担う節単位には応答を期待できる、新情報を担う名詞単位には応答があるかどうかどちらとも言えないことである。対話における話し言葉生成システムを設計

応答	名詞単位		節単位	
	旧情報	新情報	旧情報	新情報
有り	21 (73.6) 37.6	114 (119.2) 0.2	172 (204.6) 5.2	593 (502.6) 16.3
無し	79 (26.4) 104.7	48 (42.8) 0.6	106 (73.4) 14.5	90 (180.4) 45.3
テーブル χ^2 値				224.3
棄却域 (確率 = 0.05)				$7.81 < \chi^2$

表 2: 発話単位の統語カテゴリと情報の新旧が発話単位への応答の有無に与える影響

する上では、システム発話に対して応答があった方が対話相手の理解状態を推測し易い。したがって次の原則を設定する。

原則 (p3) 新情報を導入するために節単位を使え。

また、旧情報を担う単位に応答がないことは期待されるが、新情報を担う節単位に応答がないことは統計データには反することがわかる。したがって次の原則を導入する。

原則 (p4) 対話相手が旧情報を担う発話単位に回答しない場合には、計画通りに発話を継続せよ。

原則 (p5) 対話相手が新情報を担う節単位に回答しない場合には、次に何を発話すべきかについて再考せよ。

3.3 対話相手の応答に対処するための談話戦略

原則 (p5) を具体化するために、対話参加者 P_1 が新情報を担う節単位 U を発話したとき、 U に対する対話相手 P_2 の応答の有無と P_1 が後続談話を展開する方法との間の関係を分析した。 U に対して対話相手 P_2 から応答があった場合には、その応答を処理するための部分対話が終了した後 P_1 が後続談話を開始した場合のみを分析し、対話相手 P_2 が対話の統御をとって後続談話を開始した場合は分析対象から外した。発話単位 U と後続談話の間の談話関係を修辭構造理論 [8] に従ってラベル付けし、後続談話の展開の種類を次のように分類した:

Continue: U の核の記述するための後続談話が始まる。

Retain: U の内容を詳細化するか U の衛星を記述する後続談話が始まる。

Shift: その他の談話展開。

典型的な *Shift* 展開は、発話単位 U によって行為を提案し、後続談話において別の行為を提案する場合である。

分析結果を表 3 に示す。我々は、システムが新情報を担う発話単位を生成した後、応答がない場合にどうすべきかを知りたい。分析結果から、応答のない場合には

応答	Continue	Retain	Shift
有り	168 (175.2) 0.3	126 (141.5) 1.7	186 (163.4) 3.1
無し	40 (32.8) 1.6	42 (26.5) 9.0	8 (30.6) 16.7
テーブル χ^2 値			32.4
棄却域 (確率 = 0.05)			$5.99 < \chi^2$

表 3: 新情報を担う節単位への応答の有無が応答後の談話展開に与える影響

Shift展開は起こりにくいことがわかる。対話相手が応答をしない場合に Continue または Retain 展開が起きた場合は 82 例あった。この 82 例についてさらに分析を進めると、1 例を除くすべての場合において、 U の内容を詳細化した発話単位か U の衛星または核を記述する発話単位が承認されるまで Shift 展開は起こらなかった。対話相手が発話単位を承認したとは、(i) 対話相手が発話単位に対して肯定的な間投的表現 (例: 「はい」) や承認のためのメタ発話 (例: 「わかりました」) を発話した場合か、(ii) 対話相手が発話単位を確認するための発話を行い、その確認発話を処理するための部分対話が完了した場合のいずれかを言う。したがって次の原則を設定する:

原則 (p6) 対話相手が新情報を担う節単位 U に応答しないなら、対話相手が承認するまで、単位 U の内容を詳細化した発話単位か、単位 U の核または衛星を記述する発話単位を生成し続けよ。

次に、対話相手が発話単位に回答した場合の談話展開について調べた。発話単位への典型的な応答として、肯定的な間投的表現、発話単位の復唱について分析を行った。復唱には、発話単位に含まれる名詞句を完全に復唱する場合と部分的に復唱する場合がある。分析の結果、次の原則が導かれた (分析の詳細については紙面の制限のため省略する):

原則 (p7) 対話相手が承認のための肯定的な間投的表現を発話するなら、発話を計画通りに継続せよ。

原則 (p8) 対話相手が発話単位 U への応答として表現 R を発話し、 R が U に含まれる名詞句の完全な復唱なら、承認のための肯定的な間投的表現を発話せよ。

原則 (p9) 対話相手が発話単位 U への応答として表現 R_1 を発話し、 R_1 が U に含まれる名詞句 NP の部分的な復唱なら、次のいずれかを行え。

(a) 名詞句 NP の完全な復唱を行う。

(b) 名詞句 NP の部分的な復唱 R_2 を発話する。ただし

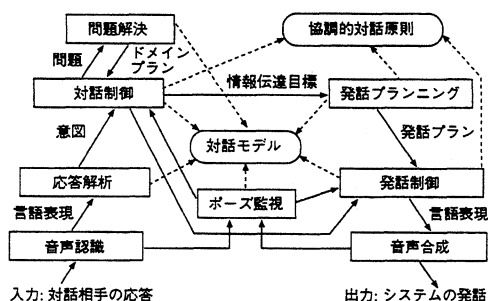


図 1: 対話における話し言葉生成システム

R_2 は R_1 と R_2 の接続が NP となるような表現である。
(c) 名詞句 NP で記述された対象の内容を詳細化する。

4 話し言葉生成システム

対話相手の応答や割り込みに対処しながら話し言葉を生成するシステムについて簡単に紹介する [4]。システム構成を図 1 に示す。このシステムは、問題解決、発話プランニング、発話制御、対話制御、応答解析、音声認識、音声合成、ボーズ監視の 8 つのモジュールから構成される。各モジュールは互いに並列に動作する。対話モデルは、対話の履歴、注視状態、ユーザーモデルなどを管理するものである [4]。協調的対話原則は、生成される談話が適切であることと対話相手からの応答に適切に対処することができることを保証する。

問題解決モジュールは与えられた問題を解き、発話プランニングモジュールは問題の解を対話相手に提案するための発話プランを立案する。発話制御モジュールは発話プランに基づいて言語表現を音声合成モジュールに送る。このシステムは、各モジュールを並行して動作させ問題を段階的に解くことにより、時間制限の下で話し言葉を生成することができる [2, 3]。

システムの発話の途中で対話相手からの応答や割り込みが発生すると、音声認識モジュールがそれを言語表現に変換し、さらに応答解析モジュールが対話相手の意図を分類する。対話相手の意図にしたがって、対話制御モジュールは発話を中断し、対話の中でのやり取りについての共通理解を保持するための付加的な発話を生成するように発話プランニングモジュールに指令する。その後、中断されていた発話は新たな文脈の下で再開される [4]。

5 協調的対話原則の効果

このシステムは計算機上に実装されており、システムとユーザとの間の対話シミュレーション実験を行うことによって、協調的対話原則が話し言葉生成において重要な役割を果たすことを確認している。システム (S) とユーザ (U) との間の対話例を図 2 に示す。システムは、小田

S: (1) 小田急線に乗っ<はい>て /
 S: (2) 愛甲石田で降りてください /
 S: (3) ここが / 小田急線の最寄りの駅になりますので /
 U: (4) はい
 S: (5) そこで降りて /
 S: (6) 北口バス乗り場というバス乗り場が
 U: (7) 北口バス乗り場
 S: (8) はいありますので /
 S: (9) そこで / 森の里青山行きというバスに
 U: (10) もり
 S: (11) の里青山行き / <はい> それに乗って /

図 2: 対話例

急線に乗って、愛甲石田で降りて、次に北口バス乗り場から森の里青山行きというバスに乗るというドメインプランを提案している。記号“/”はシステムの発話単位の区切りを示す。このような発話を実現する発話プランニング技術については、文献[3]を参照してほしい。

ユーザのアイヅチ、復唱に対してシステムは臨機応変に対処できる。これは、(i) 音声認識、対話制御、発話プランニング、発話実行などの各モジュールが並列に動作する、(ii) 原則(p1), (p2), (p3)に従ってシステムは段階的に情報をユーザに伝達するため、ユーザが応答や割り込みを行いやすい、(iii) 原則(p7), (p8), (p9)に従うことによりシステムはユーザの応答に適切な戦略で対処できるという理由による。さらに重要なことには、原則(p4), (p5), (p6)に従うことにより、ユーザの応答がなくてもシステムは談話戦略を変更することがある。図2の対話例において、システム発話(2)は新情報を担う節単位であり、この発話にユーザは応答しなかった。原則(p5), (p6)に従うことにより、システムは、その時点で計画済みの発話(6)の実行を中断して、発話(2)の衛星となる発話(3)を新たに立案し実行した。この発話(3)は、ユーザが愛甲石田で降りるという行為を採用するように動機付けるための発話である。これらの原則がないと、システムはユーザの不理解を含意する沈黙に適切に対処できない。

6 おわりに

本稿では、対話相手からの応答や割り込みに対処しながら話し言葉を漸次的に生成するシステムを示した。特に、そういったシステムがユーザとの対話に参加できるための協調的対話原則をコーパスにおける話し言葉の特性の統計的分析によって導いた。協調的対話原則の効果は対話シミュレーション実験により確認されている。

謝辞: 日頃より御指導頂く石井健一郎情報科学研究部長、有益な示唆を頂く対話理解研究グループの諸氏に感謝致

します。

参考文献

- [1] Koenraad De Smedt and Gerard Kempen. Segment grammar: A formalism for incremental sentence generation. In C. L. Paris, W. R. Swartout, and W. C. Mann, editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, pp. 329-349. Kluwer Academic Publishers, 1991.
- [2] Kohji Dohsaka and Akira Shimazu. A computational model of incremental utterance production in task-oriented dialogues. In *Proc. of COLING-96*, pp. 304-309, 1996.
- [3] 堂坂 浩二, 島津 明. タスク指向型対話における漸次的発話生成モデル. 情報処理学会論文誌, Vol. 37, No. 12, pp. 2190-2200, 1996.
- [4] 堂坂 浩二, 島津 明. 協調的な話し言葉生成. 電子情報通信学会技術研究報告 NLC96-32, pp. 9-16, 1996.
- [5] M. A. K. Halliday. *An Introduction to Functional Grammar*. Edward Arnold, 1994.
- [6] 川森 雅仁, 島津 明. 対話における発話交代の分析. 電子情報通信学会 信学技報 NLC95-73, pp. 31-38, 1996.
- [7] Gerard Kempen and Edward Hoekamp. An incremental procedural grammar for sentence formulation. *Cognitive Science*, Vol. 11, pp. 201-258, 1987.
- [8] William C. Mann and Sandra A. Thompson. Rhetorical structure theory: Towards a functional theory of text. *Text*, Vol. 8, pp. 243-281, 1988.
- [9] Johanna D. Moore and Cécile L. Paris. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, Vol. 19, pp. 651-694, 1994.
- [10] Norbert Reithinger. POPEL - a parallel and incremental natural language generation. In C. L. Paris, W. R. Swartout, and W. C. Mann, editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, pp. 179-199. Kluwer Academic Publishers, 1991.
- [11] Steve Whittaker and Phil Stenton. Cues and control in expert-client dialogues. In *Proceedings of 26th Annual Meeting of the Association for Computational Linguistics*, pp. 123-130, 1988.