

Title	音声対話システムの構成法に関する研究
Author(s)	堂坂, 浩二
Citation	
Issue Date	2004-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/947
Rights	
Description	Supervisor: 島津 明, 情報科学研究科, 博士

博 士 論 文

音声対話システムの構成法に関する研究

指導教官 島津 明 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

堂坂 浩二

2004年3月24日

要 旨

本論文は、音声対話システムの各構成部を適切に連動させることにより、システムとユーザの間に円滑な音声対話コミュニケーションを実現することを目的として、音声対話システムの構成法について論じる。

音声対話システムは、発話理解部、発話生成部、対話制御部から構成される。本論文では、システムが自然で効率的な対話を行うために構成部がどのように連動して働けばよいかという観点からシステムの構成法を論じる。第一に、ゼロ代名詞の指示対象を決定するための発話理解法を示す。提案法は、言語表現の使用に際して課せられる語用論的条件を利用して、膨大なドメイン知識に頼ることなく、ゼロ代名詞の指示対象を決定する。発話理解部と発話生成部が文脈情報を共有して連動することにより、提案法は、ユーザ発話理解だけでなく、文脈の下での適切なシステム発話の生成にも役立てることができ、自然な対話の実現が可能となる。

第二に、システム発話途中のユーザ応答(アイツチ、割り込み発話)に対処しながら自然な発話を生成するための漸次的発話生成法を示す。話し言葉対話では書き言葉の文よりも小さな発話単位でやり取りが行われる。したがって、書き言葉の文を単位として発話を生成する従来法では、ユーザ応答に即座に対処することは難しい。提案法では、発話生成部は話し言葉特有の小さな発話単位で漸次的に発話を生成する。人同士の対話の分析から導いた協調的対話原則に基づいて、発話生成部が対話制御部、発話理解部と連動することにより、ユーザ応答に臨機応変に対処しながら自然な発話を生成することが可能となる。

第三に、できるだけ短い対話でユーザが必要とする情報を伝達し、効率的な対話を実現するための対話制御法として、デュアルコスト法と試行型デュアルコスト法を示す。提案法は、音声認識精度に依存する確認コストとシステムのデータベース内容に依存する情報伝達コストという2つのコストの和として対話の長さを推定する。対話制御部が発話理解部、発話生成部と連動して、対話全体の長さを最小とするように対話の各時点におけるシステム行動を選択することにより、提案法は従来法よりも効率的な対話を実現することができる。シミュレーション対話実験と被験者対話実験の結果により提案法の効果を実証する。

Abstract

In this thesis, we discuss an architecture for spoken dialogue systems that achieves smooth conversation between a system and its users through appropriate cooperation among system components.

A spoken dialogue system is composed of utterance understanding, utterance generation, and dialogue control components. We discuss a system architecture with emphasis on how the components should cooperate with each other to achieve a natural and efficient dialogue. First, we present an utterance understanding method based on pragmatic constraints for identifying the referents of zero-pronouns. Through cooperation between the utterance understanding and generation components sharing contextual information, the proposed method not only understands user utterances but also serves for generating relevant system utterances under the context, and it enables a system to achieve a natural dialogue.

Second, we present an incremental utterance production method that allows a system to generate utterances while handling user responses. In spoken dialogue, dialogue participants use utterance units that are smaller than sentences in written language. Therefore, conventional methods utilizing sentences as utterance units cannot rapidly handle user responses during system utterances. The proposed method generates utterances incrementally in small utterance units specific to spoken dialogues. The utterance generation component cooperates with the dialogue control and utterance understanding components based on collaborative principles derived from an analysis of human-human dialogues. Through this cooperation, a system can generate natural utterances while handling user responses.

Third, we present dialogue control methods called dual-cost and trial dual-cost methods for a system to convey the desired information to the user in as short a dialogue as possible. The dialogue length is estimated as the sum of the confirmation cost depending on the speech recognition rate and the information transfer cost depending on the system database content. By choosing a system action at each point of a dialogue so as to minimize the total dialogue length through cooperation among the dialogue control, utterance understanding, and utterance generation components, the proposed methods achieve a more efficient dialogue than conventional ones. The results of experiments by dialogue simulation and with human users prove the effectiveness of the proposed methods.

目次

1	序論	1
1.1	研究の背景	1
1.1.1	発話理解における問題	2
1.1.2	発話生成における問題	3
1.1.3	対話制御における問題	4
1.2	研究の目的	4
1.2.1	語用論的条件の充足に基づく発話理解	5
1.2.2	漸次的発話生成	5
1.2.3	効率的対話制御	7
1.3	論文の構成	8
2	音声対話システム	9
2.1	はじめに	9
2.2	主導権混在型音声対話システム	10
2.2.1	対話の主導権による音声対話システムの分類	10
2.2.2	主導権混在型システムとユーザの対話	11
2.3	本研究の位置付け	14
2.4	本研究におけるシステム構成	16
2.5	まとめ	21
3	語用論的条件の充足に基づく発話理解	22
3.1	はじめに	22
3.2	語用論的条件の充足に基づく発話理解法の概観	26
3.3	言語表現の使用と語用論的条件	29
3.3.1	待遇表現の使用と社会関係	29

3.3.2	受給表現の使用と話し手の視点	30
3.3.3	文末形式と情報のなわばり	30
3.4	文脈	33
3.5	語用論的条件の充足に基づく発話理解プロセス	37
3.5.1	単一化	37
3.5.2	理解プロセス	37
3.5.3	理解結果の選択	38
3.5.4	ゼロ代名詞の指示対象の決定例	40
3.6	評価	44
3.7	まとめ	48
4	漸次的発話生成	50
4.1	はじめに	50
4.2	対話データに基づく協調的対話原則の分析	54
4.2.1	発話単位の分析	55
4.2.2	談話関係の分析	57
4.2.3	対話相手の応答に対処するための談話戦略の分析	59
4.3	漸次的発話生成法の概観	65
4.4	協調的対話原則と対話状態	68
4.5	問題解決	69
4.6	発話プランニング	71
4.6.1	対話目標	72
4.6.2	対話行為	72
4.6.3	階層的プランニング技法による発話プランニングの実現法	73
4.6.4	発話実行	80
4.6.5	発話プランニングと発話実行の並行進行による漸次的発話生成	81
4.6.6	対話相手からの応答に対処するための対話制御	83
4.7	考察	85
4.8	おわりに	90

5	効率的対話制御	93
5.1	はじめに	93
5.2	音声対話システムにおける対話制御	98
5.3	デュアルコスト法	101
5.3.1	概観	101
5.3.2	動作手順	103
5.3.3	ユーザ問い合わせタイプの確率分布	104
5.3.4	対話プランの生成	105
5.3.5	対話コストの計算	107
5.4	試行型デュアルコスト法	109
5.5	シミュレーション対話実験による評価	111
5.5.1	シミュレーション対話実験	111
5.5.2	結果と考察	115
5.6	被験者対話実験による評価	120
5.6.1	予備実験	121
5.6.2	評価実験	121
5.6.3	結果と考察	122
5.7	おわりに	125
6	結論	127
6.1	本研究の要約	127
6.2	今後の展望	129
	謝辞	131
	研究業績	143

目 次

2.1	主導権混在型システムとユーザの対話の流れ	12
2.2	音声対話システムの構成	16
3.1	ゼロ代名詞の使用例	23
3.2	語用論的条件の充足に基づく発話理解法のシステム構成	26
3.3	情報のなわばりについての対話例	33
3.4	社会関係についての制約	34
3.5	話し手の視点と情報のなわばりについての制約	36
3.6	理解プロセス	39
3.7	提案法が指示対象を決定できなかったゼロ代名詞の例	46
4.1	書き起こされた対話からの抜粋	54
4.2	図 4.1 に示した談話における談話関係	58
4.3	漸次的発話生成のシステム構成	66
4.4	路線図の例	70
4.5	Sequence 関係に基づく発話プランニングのための諸定義	75
4.6	Elaboration 関係に基づく発話プランニングのための諸定義	76
4.7	Circumstance 関係に基づく発話プランニングのための諸定義	78
4.8	Motivate 関係に基づく発話プランニングのための諸定義	80
4.9	システムが生成した談話 (d4.5)	86
4.10	システムが生成した談話 (d4.6)	86
4.11	ユーザのアイツチや復唱へ対処しながらの対話 (d4.7)	88
4.12	ユーザの割り込み発話に対処しながらの対話 (d4.8)	90
5.1	不必要な確認を伴う非効率な対話の例	94
5.2	不必要な確認を避けた効率的な対話の例	95

5.3	試行型の情報伝達行為の例	101
5.4	デュアルコスト法のシステム構成	102
5.5	デュアルコスト法の動作手順	103
5.6	一括確認法と個別確認法の動作手順	112
5.7	無確認法の動作手順	113
5.8	模擬ユーザの振る舞いを決定するルール	114
5.9	警報問い合わせの場合における属性認識精度に応じた対話の長さの 平均	115
5.10	天気問い合わせの場合における属性認識精度に応じた対話の長さの 平均	116
5.11	気温問い合わせの場合における属性認識精度に応じた対話の長さの 平均	116
5.12	降水確率問い合わせの場合における属性認識精度に応じた対話の長 さの平均	117
5.13	各問い合わせタイプを無作為に発生させた場合における属性認識精度 に応じた対話の長さの平均	117
5.14	被験者対話実験における各対話制御法による対話の長さの平均 . . .	123

表 目 次

3.1	語用論的条件と文脈を表現するための事態	28
3.2	発語内行為タイプ	31
3.3	文のタイプと情報のなわばり	32
3.4	語用論的条件の充足に基づく発話理解法の評価結果	45
4.1	名詞単位と節単位に関して, 1 単位中に現れる名詞句の数の頻度分布	56
4.2	談話関係の頻度分布	58
4.3	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	60
4.4	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	63
5.1	天気情報案内タスクにおけるユーザ問い合わせタイプと属性の関係	99
5.2	被験者対話実験における各対話制御法による対話の長さの平均 . . .	124

第 1 章

序論

1.1 研究の背景

人と日常の言葉で会話できるコンピュータの実現は、半世紀に渡って計算機科学における挑戦的な研究課題の一つであり続けてきた [82]。「人のように会話するコンピュータ」という壮大な目標は、研究者の関心を引き続けてきただけでなく、サイエンスフィクション「2001 年宇宙の旅」に登場するコンピュータ HAL に代表されるように、多くの人の夢をかきたててきた。

近年の音声情報処理、自然言語処理の発展と、コンピュータの計算能力の向上にともなって、人同士が日常行っている自然な音声会話を人とコンピュータの間に実現しようとする研究が盛んに行われるようになった。ユーザとの音声対話を通して、情報案内、意思決定支援といった決められたタスクを実行するシステムは音声対話システムと呼ばれ、これまで様々なタスクの音声対話システムが開発されてきている [3, 28, 31, 64, 69, 91]。

音声対話システムは、データベース管理などのアプリケーション・プログラムに対して音声対話インタフェースを提供するものとみなすことができる。音声対話インタフェースには次のような利点がある [61, pp. 1–2]。

- 音声対話は、人が日常用いているコミュニケーション手段であり、音声対話インタフェースはユーザにとって親しみやすいインタフェースとなることが期待される。

- 音声対話インタフェースは、手や目を占有しないので、ユーザは他の作業に従事しながらでも使用可能である。
- ユーザは、特別な訓練を行うことなく、音声対話インタフェースを通して情報を生成することができる。
- キーボードやタイプライターを使ったテキスト入力の会話インタフェースに比べると、音声対話インタフェースはユーザとシステムの間で高速な情報交換が可能である。

このように音声対話インタフェースは多くの利点を持ち、システムとユーザの間に円滑な音声対話コミュニケーションを実現できれば、使い勝手の良いインタフェースとなることが期待される。

一般に、音声対話システムは、入力されたユーザ音声からユーザの発話内容を組み立てる発話理解部、対話の各時点において適切なシステムの行動を決定する対話制御部、対話制御部の決定にしたがって言語表現を生成・音声出力する発話生成部から成る。本研究では、音声対話システムとユーザの間に円滑な音声対話コミュニケーションを実現するために、システムの構成部がどのように連動して働くべきかという観点から、音声対話システムの発話理解、発話生成、対話制御の各機能に関して、以下に述べる問題に着目する。

1.1.1 発話理解における問題

本研究は、発話理解の中でも、ユーザ発話内容を対話の文脈と関連付けて理解する談話理解に関心がある。日常の音声会話においては、代名詞や省略表現などの文脈依存表現が頻繁に使用される。これらの文脈依存表現は、それ自身は簡潔な表現であるにもかかわらず、文脈によって多様な内容を伝達できるという意味で、効率的であると呼ばれる [5]。音声対話システムが文脈依存表現の内容を文脈の下で適切に理解することができなければ、ユーザは効率的ではない冗長な言語表現を使わざるをえなくなり、音声対話の円滑さが阻害される。

文脈依存表現は、音声対話だけでなく、書き言葉によるテキストにおいても現れるが、特に日本語の話し言葉に特有の問題として、発話の中での対話登場人物

への指示が、多くの場合にゼロ代名詞化されるということがある [39]。ゼロ代名詞とは表層では形をもたない代名詞である。円滑な音声対話を実現するためには、ユーザ発話に現れるゼロ代名詞を正しく理解する必要がある。本研究では、日本語対話において対話登場人物を指示するゼロ代名詞の指示対象を決定するという問題に着目する。さらに、ユーザ発話は文脈の下で理解されるだけでなく、ユーザ発話によって新規な文脈情報が導入され、導入された文脈情報に基づいて、後続対話におけるユーザ発話の理解とシステム発話の生成が行われる必要があることに着目する。

1.1.2 発話生成における問題

本研究は、発話生成の中でも、一連の発話を首尾一貫とした談話として生成する談話生成に関心がある。自然言語処理において、言語生成の研究は書き言葉による文の生成から始まり、その後、書き言葉による談話の生成に発展していった。しかし、音声対話における話し言葉による談話の生成は、書き言葉による談話の生成にはない特有の問題が存在する。

音声対話は複数の対話参加者による実時間の協同活動である。音声対話システムは、発話を生成するとき、ユーザに対して一方的に情報を提示するだけでは不十分であり、システム発話途中のユーザからのアイツチや割り込み発話に応じながら発話を生成する必要がある。このとき、システムは発話する前にいくらでも時間をかけて発話内容について熟考して良いわけではなく、発話すべきときに即座に発話を開始することが望ましい。したがって、システムは、発話途中のユーザのアイツチや割り込み発話に即座に応じて、臨機応変に話の進め方を変更する必要がある。このことによって、システムはユーザの理解状態を確認しながらユーザに必要な情報を伝達し、ユーザは望む情報を迅速にシステムから聞き出すことが可能となり、システムとユーザの円滑な対話を実現できる。

以上の観点から、本研究では、音声対話システムがシステム発話途中のユーザのアイツチや割り込みに即座に対処し、ユーザの意図に応じて臨機応変に話の流れを変更しながら、自然な発話を生成するという問題に着目する。

1.1.3 対話制御における問題

対話制御は，人工知能の分野において従来より研究されてきた自律エージェントの行動プラン立案とプランに基づく行動実行という枠組み [76] の中で捉えることができる．本研究では，音声対話システムが音声認識誤りの下でタスクを遂行するために，いかにシステムの行動を選択するかという音声対話に固有の問題に関心がある．

音声対話システムでは，音声認識誤りの下でユーザ問い合わせ内容を把握するために，システムが認識したユーザ問い合わせ内容をユーザに対して確認するということがよく行われる．ユーザ問い合わせ内容の把握することを目的として行われるシステムとユーザのやり取りは確認対話と呼ばれる．確認対話において，システムは認識した問い合わせ内容が正しいかどうかをユーザに対して確認する．ユーザはシステムの確認に対して訂正発話を行うこともできるし，肯定的表現 (例:「はい」) を発話することによってシステムの確認を承認することもできる．システムは，確認対話を終えた後，把握した問い合わせ内容にしたがって，ユーザが必要とする情報を伝達するために応答を生成する．

確認対話は，音声認識誤りの下でタスクを遂行するために有効ではあるが，過度に長い確認対話は対話の円滑な流れを損なう．したがって， unnecessary 確認は避けることが望ましい．そこで，本研究では， unnecessary 確認を避けて，できるだけ短い対話でユーザが必要とする情報を伝達することを可能とする対話制御に着目する．

1.2 研究の目的

本研究の目的は，音声対話システムの各構成部を適切に連動させることにより，システムとユーザの間の円滑な音声対話コミュニケーションを実現することにある．この目的の下に 1.1 節で取り上げた問題に着目し，以下に述べる 3 つの研究テーマに取り組む．本論文では，各テーマにおいて，システムとユーザの間の円滑な対話を実現するために，システムの構成部がどのように連動して働けばよいかという観点からシステムの構成法を論じていく．

1.2.1 語用論的条件の充足に基づく発話理解

文脈依存表現として、日本語対話において対話登場人物を指示するゼロ代名詞を取り上げ、語用論的条件の充足に基づいてゼロ代名詞の指示対象を決定する方法を提案する [19, 21] .

文脈依存表現理解に関しては、ユーザの意図 [1, 8, 51, 52] , 注視状態 [7, 41, 86] という文脈情報を利用する方法が従来より提案されてきた。これらの従来法は、ドメインのイベントや行為の因果関係に関する知識の存在を前提としており、そういった知識を書き下すことは多大な労力を要するという問題があった。本研究では、対話登場人物間の社会関係 [59] , 話し手の視点 [48] , 情報のなわばり [42] という文脈情報を利用して、日本語対話文中において対話登場人物を指示するゼロ代名詞の指示対象を決定する発話理解法を提案する。提案法は、ドメインの行為やイベントの因果関係に関する知識に頼ることなく、ゼロ代名詞の指示対象を決定できるという利点をもつ。

提案法は、ユーザ発話が入力されると、待遇表現、受給表現などの言語表現を適切に使用するために課せられる語用論的条件を抽出し、抽出した語用論的条件を文脈の下で充足させることにより、ゼロ代名詞の指示対象を決定する。文脈には、対話登場人物間の社会関係、話し手の視点、情報のなわばりに関する制約が書かれている。

提案法では、ユーザ発話中のゼロ代名詞の指示対象が決定されるだけでなく、ユーザ発話によって伝達される新たな文脈情報を導き出し、文脈に導入される。発話理解部と発話生成部が文脈情報を共有して連動することにより、ユーザ発話により新規に導入された文脈情報を後続の対話におけるユーザ発話理解やシステム発話生成に役立てることが可能となる。

1.2.2 漸次的発話生成

システムがユーザ問い合わせ内容に応じてユーザが必要とする情報を伝達するときに、ユーザからのアイツチや割り込み発話に即座に対処しながら、一連の発話を自然な談話として漸次的に生成する方法を提案する [22, 23, 24] .

従来の言語生成法は書き言葉における文を単位として発話を生成していた [38,

55, 57, 58, 60, 67] . しかし , 音声対話では書き言葉における文よりも小さな単位で発話が生成されるので , システムが書き言葉における文を発話している途中であってもユーザのアイツチや割り込み発話が起きる . システムが文を単位として発話を生成していると , その途中に起きたユーザのアイツチや割り込み発話に即座に対処できないだけでなく , アイツチや割り込み発話が文のどの構成素に対して成されたのか判別できないため , システムはユーザのアイツチや割り込み発話に適切に対処できない . したがって , 従来の書き言葉の文を単位する言語生成法では , ユーザのアイツチや割り込み発話に即座に対処して , 臨機応変に話の進め方を変更しながら , 発話を生成することは困難である .

本研究では , 話し言葉対話における発話の特徴についての知見を得るために , 人同士の音声対話の書き起こしデータを使って , 発話単位の大きさや対話相手の応答に対処するための談話戦略について分析し , システムがユーザの応答に対処しながら発話を生成するために必要となる協調的対話原則を導く .

この分析結果を踏まえ , システムが協調的対話原則にしたがって話し言葉特有の小さな発話単位で発話を漸次的に生成する方法を示す . 漸次的な発話生成とは , 発話内容を組み立てながら発話を生成し , 発話を行いながら次の発話内容を決定していくことを言う . 小さな発話単位で段階的に情報を伝達することにより , システム発話途中のユーザからの応答に即座に対処することが可能となる . 加えて , 発話内容を組み立てながら発話を生成することにより , 発話内容や談話構造が最終的に決定されるのを待たずに発話を開始できるので , システムが発話すべきときに即座に発話を開始することが容易となる .

提案法では , 小さな発話単位で漸次的に発話を生成する発話生成部が , 協調的対話原則にしたがって , 発話理解部 , 対話制御部と適切に連動することにより , システム発話途中のユーザのアイツチ , 割り込み発話に即座に対処し , ユーザ意図に応じて臨機応変に話の流れを変更する . 結果として , システムは , 対話の中でやり取りされる情報についてユーザとの間で共有理解を維持しながら , 自然な発話を生成することが可能となる .

1.2.3 効率的対話制御

音声認識精度とシステムのデータベース内容にしたがって、システムができるだけ短い対話でユーザが必要とする情報を伝達することを可能とする対話制御法を提案する [25, 26, 27] .

1.1.3 節で説明したように、音声対話システムでは、音声認識誤りに対処するために、認識した問い合わせ内容をユーザに確認することが行われる。効率的に確認対話を実施することを目指した対話制御方法が従来より提案されているが [11, 50, 53, 66, 75] , 従来法は、確認のためのやり取りの長さを削減することに注目し、確認後のシステム応答の長さを考慮しないため、不必要な確認を回避できない場合があった。

このことを天気情報案内システムを例にとって説明する。どこにも警報が発表されていないというデータベース内容の下で、ユーザが「東京に大雨警報が発表されているかどうか」と問い合わせたと認識された場合を想定する。このとき、システムは警報であることを確認しさえすれば、「どこにも警報は発表されていません」と応答することができるので、場所や警報の種類といった項目の確認は不必要であり、回避することが望ましい。これらの項目を確認したとすると、確認の途中で音声認識誤りが生じた場合、ユーザが必要とする情報の伝達に貢献しない不必要なやり取りが延々と繰り返されることになる。

データベースの内容が多くのある場所に警報が発表されているというものであれば、場所の確認は必要となる。場所の確認を省略すると、システムは警報が発表されているすべての場所についての情報をユーザに伝達することになり、対話の長さがかえって長くなるからである。このように、確認が必要かどうかを対話制御部が決定するためには、その時点のシステム理解状態を参照するだけでなく、確認後のシステム応答の長さを発話生成部に問い合わせることが必要となる。

そこで、本研究では、確認のためのやり取りと確認後のシステム応答を含めた対話全体の長さを最小にするように対話各時点のシステム行動を選択する対話制御法として、デュアルコスト法、試行型デュアルコスト法を提案する。提案法では、確認を完了するまでにやり取りされる自立語の数の期待値 (確認コスト) と確認後のシステム応答に含まれる自立語の数の期待値 (情報伝達コスト) の和として対話の長さを推定する。確認コストは音声認識精度に依存し、情報伝達コストは

データベース内容に依存する．確認コストと情報伝達コストは相反する関係にある．この２種類のコストの和を最小にするように対話各時点のシステム行動は選択される．

提案法では，対話制御部は，システム理解状態を随時更新する発話理解部と，予測されるシステム応答の長さを計算する発話生成部と連動しながら，コストの和が最小となるような対話を導くシステム行動を選択する．結果として，従来法で回避できなかった不必要な確認を回避し，効率的な対話を実現することが可能となる．

1.3 論文の構成

本論文の構成は次の通りである．第 2 章では，本研究で焦点を当てる主導権混在型の音声対話システムについて論じ，本研究の位置付けを述べる．また，各研究テーマについて，システム構成部の機能と構成部間の連動について概略を述べる．第 3 章では，語用論的条件の充足に基づいて，対話登場人物を指示する日本語ゼロ代名詞の指示対象を決定する発話理解法を示し，提案法の評価結果について論じる．第 4 章では，人同士の対話の書き起こしデータの分析により，話し言葉特有の発話単位，談話構造と，システムがユーザからの応答に対処しながら発話を生成するために必要となる協調的対話原則を導く．次に，協調的対話原則にしたがって，ユーザのアイツチや割り込み発話に対処しながら自然な発話を生成することを可能とする漸次的発話生成法を示し，提案法の有効性について考察する．第 5 章では，音声認識精度とシステムのデータベース内容に基づいて，できるだけ短い対話でユーザが必要とする情報を伝達することを可能とする対話制御法として，デュアルコスト法と試行型デュアルコスト法を示し，提案法の有効性を評価するために行った対話実験の結果について論じる．第 6 章では本研究の要約と今後の展望について述べる．

第 2 章

音声対話システム

2.1 はじめに

音声対話システムは，ユーザとの音声対話を通して，情報案内，意思決定支援といった決められたタスクを実行するシステムである．本研究は，音声対話システムの各構成部を適切に連動させることにより，システムとユーザの間の円滑な音声対話コミュニケーションを実現することを目的とし，具体的な研究テーマとして，語用論的条件の充足に基づく発話理解，漸次的発話生成，効率的対話制御という 3 つのテーマに取り組む．

音声対話システムとユーザは，互いに補完的な知識や情報を有しており，対話を通して互い協力しながら，協同で一つのタスクを遂行するものと考えられる．たとえば，ユーザが対話を通して達成しようとするユーザ意図について，ユーザはあらかじめ知っているが，システムは前もっては知らない．一方，タスクを達成するためにどういった情報を伝える必要があるかについて，システムは前もって知っているが，ユーザは必ずしも知っているとは限らない．このように互いに補完的な知識や情報を有するシステムとユーザが，円滑な対話を通して一つのタスクを遂行するためには，対話の局面によってシステムとユーザが対話の主導権 [44, 87, 88] を柔軟に交代しながら，システムとユーザの間で適切な分業を実現することが重要である．

以上の観点から，本研究は，システムとユーザが対話の主導権 [44, 87, 88] を自由に交代しながら対話を進めることができる主導権混在型の音声対話システムに焦

点をあて、円滑な音声対話を実現するために必要なシステム構成について論じる。

以下においては、まず、対話の主導権という観点から主導権混在型の音声対話システムを定義づける。次に、主導権混在型の音声対話システムにおける本研究の位置付けを述べる。続いて、音声対話システムの構成について概観し、各研究テーマにおける構成部の機能と構成部間の連動について概略を述べる。

2.2 主導権混在型音声対話システム

2.2.1 対話の主導権による音声対話システムの分類

音声対話の各時点において、対話の進め方を主導する対話参加者は対話の主導権をもつと言われる [44, 87, 88]。以下に示すように、いずれの対話参加者が主導権をもつかは発話タイプによって判別できる [87]。

発話タイプによる主導権の位置の判定 [87]

発話タイプ	主導権をもつ対話参加者
陳述	疑問への答えではない場合は話し手 疑問への答えの場合は聞き手
疑問	話し手
命令	話し手
アイツチ、復唱	聞き手

音声対話システムとユーザは、互いに補完的な知識、情報を有しており、必要に応じて対話の主導権を交代しながら対話を進めることにより、円滑な対話を通して協同でタスクを遂行することができると考えられる。

音声対話システムは、対話の主導権という観点から、ユーザ主導のシステム、システム主導のシステム、主導権混在型のシステムに分類することができる。ユーザ主導のシステムとは、主としてユーザが対話の主導権をもっているシステムであり、システムはユーザからの指示や質問に対して受動的に応答することしかできない。たとえば、音声によって機器を操作する音声コマンドシステムは、システムが主導的にユーザへ働きかける機能が組み込まれていない限り、ユーザ主導

のシステムと考えることができる。ユーザ主導のシステムでは、システムはユーザの発話に曖昧さや誤りがあっても、システムが主導して曖昧さや誤りを解決することができず、円滑な対話を実現することは難しい。

システム主導のシステムとは、主としてシステムが対話の主導権をもつシステムであり、システムからの指示や質問に対してユーザが受動的に応答することによって対話は進行する。システム主導のシステムでは、対話遷移のシナリオがあらかじめ決められており、そのシナリオに沿って、システムはユーザに指示や質問を行い、システムの指示や質問に対するユーザの応答の範囲もシナリオによって決められている [14, 29]。システム主導のシステムは、タスク構造が単純で、タスク全体が独立な部分タスクへ分割できる場合には、効率的な対話を実施できることがあるが、そうでない場合には、対話の途中で後戻りが生じる可能性が高く、円滑な対話を実現することは難しい。

主導権混在型のシステムとは、対話の途中でシステムとユーザの間で自由に主導権が交代することを許すシステムである。主導権混在型のシステムは、ユーザが主導して対話を進めているときに、システムは必要ならば主導権をとって、ユーザ発話の曖昧さや誤りを解決することができる。また、システムが主導して対話を進めているときに、対話の進め方が不適切であることにユーザが気づけば、ユーザは割り込んで主導権をとり、対話の進め方を変更することができる。このように、主導権混在型のシステムはユーザ主導やシステム主導のシステムよりも円滑な対話を実現することができる。現在、多くの主導権混在型の音声対話システムが構築されてきている [3, 28, 31, 64, 69, 91]。本研究では主導権混在型の対話システムに焦点をあてる。

2.2.2 主導権混在型システムとユーザの対話

主導権混在型の音声対話システムとユーザの対話は、ユーザ問い合わせの内容を把握するためのユーザ問い合わせ把握フェーズと、把握した問い合わせ内容に基づいてシステムがユーザに対応するシステム対応フェーズの2つのフェーズの間を移行しながら進行する。ユーザがシステムデータベース内容についての問い合わせを行い、ユーザ問い合わせ内容に応じてユーザが必要とする情報を伝達す

対話 (d2.1)

- (u2.1.1) ユーザ: 神奈川県は今後一週間の天気を教えてください。
{ ユーザ問い合わせ把握フェーズの開始 }
- (u2.1.2) システム: 天気についてですか? { 確認発話 }
- (u2.1.3) ユーザ: 神奈川県について知りたいんですが
{ システムの確認意図から外れた発話 }
- (u2.1.4) システム: いつですか? { 情報要求発話 }
- (u2.1.5) ユーザ: 今後一週間の天気です { 情報要求への応答 }
- (u2.1.6) システム: 神奈川県の日明日の天気ですか? { 確認発話 }
- (u2.1.7) ユーザ: 今後一週間の天気です { 訂正発話 }
- (u2.1.8) システム: 神奈川県は今後一週間の天気ですか? { 確認発話 }
- (u2.1.9) ユーザ: はい { 承認発話 }
- (u2.1.10) システム: 神奈川県のは今日は晴れ, 明日は雨, 木曜日は
{ システム対応フェーズの開始 }
- (u2.1.11) ユーザ: 土曜日は? { 割り込み発話 }
- (u2.1.12) システム: 土曜日は曇りの予報です { 対話の進め方を変更 }

図 2.1: 主導権混在型システムとユーザの対話の流れ

るという情報案内タスクの音声対話システムの場合，システム応対フェーズをシステム情報伝達フェーズとも呼ぶ．

ユーザ問い合わせ把握フェーズにおけるシステムとユーザのやり取りは確認対話と呼ばれる．システムは音声認識誤りの下で問い合わせ内容を把握するために確認対話を行う．確認対話では，システムは，確認発話（例：「お尋ねの場所は神奈川県ですか？」）によって，認識した問い合わせ内容をユーザに対し確認し，情報要求発話（例：「いつですか？」，「どこですか？」）によって，問い合わせ内容を把握するために必要な情報をユーザから引き出す．ユーザは，システムの確認発話に対して，確認内容が正しければ承認発話（例：「はい」）を行い，確認内容が誤っていれば訂正発話を行うことができる．ユーザによって承認された問い合わせ内容は承認済みであると呼ぶ．ユーザはシステムの情報要求発話に対して必要な情報をシステムに伝達する．ユーザは常にシステムの確認・情報要求の意図にしたがって応答する必要はなく，システムの意図を拒否したり，システムの意図から外れた発話を行うこともできる．確認対話はシステムが音声認識の下でユーザ問い合わせ内容を的確に把握することに役立つ．

ユーザ問い合わせ把握フェーズで問い合わせ内容を把握した後，システムはシステム応対フェーズに移行する．確認対話の終了時点でシステムが問い合わせ内容として何を正しいと信じるかによって，2つのシステム対話戦略を考えることができる．一つはユーザが承認した内容だけを正しいとする確定型の対話戦略であり，もう一つはユーザが承認していない内容も正しいとする試行型の対話戦略である．確定型の対話戦略は問い合わせ内容に適合した応対を行う確実性が増すが，確認対話が長くなる傾向がある．試行型の対話戦略は確認対話の長さを削減することができるが，問い合わせ内容に適合した応対を行う確実性が減少する．

例として，図 2.1 に天気情報案内を行う音声対話システムとユーザの対話を示す．(u2.1.1) から (u2.1.9) までがユーザ問い合わせ把握フェーズ（確認対話）であり，(u2.1.10) 以降がシステム応対フェーズ（システム情報伝達フェーズ）である．ユーザ問い合わせ把握フェーズにおいて，システムは，(u2.1.2)，(u2.1.6)，(u2.1.8) で確認発話を行い，(u2.1.4) で情報要求発話を行っている．ユーザは，(u2.1.3) ではシステムの確認の意図に反して，システムから主導権を取って，対話の進め方を変更している．ユーザは，(u2.1.5) ではシステムの情報要求発話にしたがって必要な情報

を提示し, (u2.1.7) ではシステムの確認に対して訂正発話を行い, (u2.1.9) では承認発話を行っている。システム応対フェーズにおいて, システムは, (u2.1.10) で主導権をとって, ユーザが必要とする情報の伝達を開始している。ユーザは, (u2.1.11) でシステムに割り込んで主導権を奪い, 対話の進め方を変更している。このように, それぞれの対話フェーズの中で主導権の自由な交代が起きている。

2.3 本研究の位置付け

システムとユーザが円滑に対話を進めるためには, システムとユーザが自由に主導権を取り合ったり, 主導権を相手に譲ったりといった柔軟な主導権交代を実現することが必要である [24, 64]。本研究は, 主導権変動型の音声対話システムにおける円滑な音声対話コミュニケーションの実現を目的とする研究と位置付けることができる。具体的な研究テーマとしては, 語用論的条件の充足に基づく発話理解, 漸次的発話生成, 効率的対話制御という3つの研究テーマを取り上げる。各研究テーマは, 主導権混在型システムの異なる側面に焦点をあてたものとして位置付けることができる。

語用論的条件の充足に基づく発話理解は, ユーザ問い合わせ把握フェーズにおいて, システムがユーザの発話内容に対話の文脈と関連付けて理解するときに, 対話登場人物を指示するゼロ代名詞の指示対象を決定することに焦点をあてる。第3章で述べるように, 文脈に基づいてユーザ発話内容を理解する従来法としては, プラン認識に基づいて文脈依存表現を理解する方法 [1, 8, 51, 52] や, センタリング理論に基づいて代名詞の指示対象を決定する方法 [7, 41, 86] があるが, これらの従来法は, 行為やイベントに関する膨大な知識を前提とする問題がある。本研究では, 待遇表現などの特定の言語表現を適切に使用するための語用論的条件を利用することによって, 行為やイベントに関する膨大な知識に頼ることなく, 対話登場人物を指示するゼロ代名詞の指示対象を決定できる。提案法は, ユーザ発話の内容を理解するだけでなく, ユーザ発話によって導入される文脈情報を取り出すことができる。発話理解部と発話生成部が文脈情報を共有して連動することにより, ユーザ発話により導入された文脈情報を後続の対話におけるユーザ発話理解やシステム発話生成に役立てることができる。

漸次的発話生成は，システム対応フェーズにおいて，柔軟な主導権交代を実現するために，システムが発話途中のユーザのアイツチや割り込み発話に即座に対処し，臨機応変に話の流れを変更しながら，自然な発話を生成することに焦点をあてる．第4章で述べるように，従来の談話生成の研究 [38, 55, 57, 58, 60, 67] は書き言葉の文を単位として発話を生成することに着目してきた．しかし，話し言葉による対話では，書き言葉における文よりも小さな単位で発話が生成されることが知られており [16, 22, 35, 49, 68, 70]，従来の談話生成法では，話し言葉特有の小さな発話単位に対するユーザの応答に即座に対処しながら，発話を生成することは難しい．本研究では，話し言葉特有の小さな発話単位を使って，発話内容の組み立てと発話の実行を並行して進める漸次的発話生成法を提案する．提案法は，話し言葉特有の小さな発話単位で段階的に情報を伝達するので，ユーザのアイツチや割り込みで即座に対処できる．加えて，発話内容が最終的に決定される前に発話を開始できるので，システムが発話すべきときに即座に発話を開始することが容易である．小さな発話単位を使って漸次的に発話を生成する発話生成部が，対話制御部，発話理解部と適切に連動することにより，ユーザのアイツチや割り込み発話に即座に対処して，話の流れを臨機応変に変更しながら，自然な発話を生成することが可能となる．このようにして，提案法は，主導権混在型システムのシステム対応フェーズにおいて，柔軟な主導権交代を実現することができる．

効率的対話制御は，対話の全体を通してシステムとユーザが主導権を交代しながら会話するときに，できるだけ短い対話でユーザが必要とする情報を伝達し，効率的な対話を実現するために，対話の各時点における適切なシステム行動を選択することに焦点を当てる．第5章で述べるように，従来の対話制御法 [11, 50, 53, 66, 75] は，ユーザ問い合わせ把握フェーズにおける確認対話の長さを削減することに注目し，システム対応フェーズにおけるシステム応答の長さを考慮していないために， unnecessary 確認を行い，対話が不必要に長くなってしまう場合があった．本研究では，確認対話とシステム応答を合わせた対話全体の長さを最小にするように，対話の各時点におけるシステム行動を選択する対話制御法としてデュアルコスト法，試行型デュアルコスト法を提案する．提案法は，確認対話の長さの期待値（確認コスト）と確認後のシステム応答の長さの期待値（情報伝達コスト）の和として対話の長さを推定する．確認コストと情報伝達コストは相反する関係にある．こ

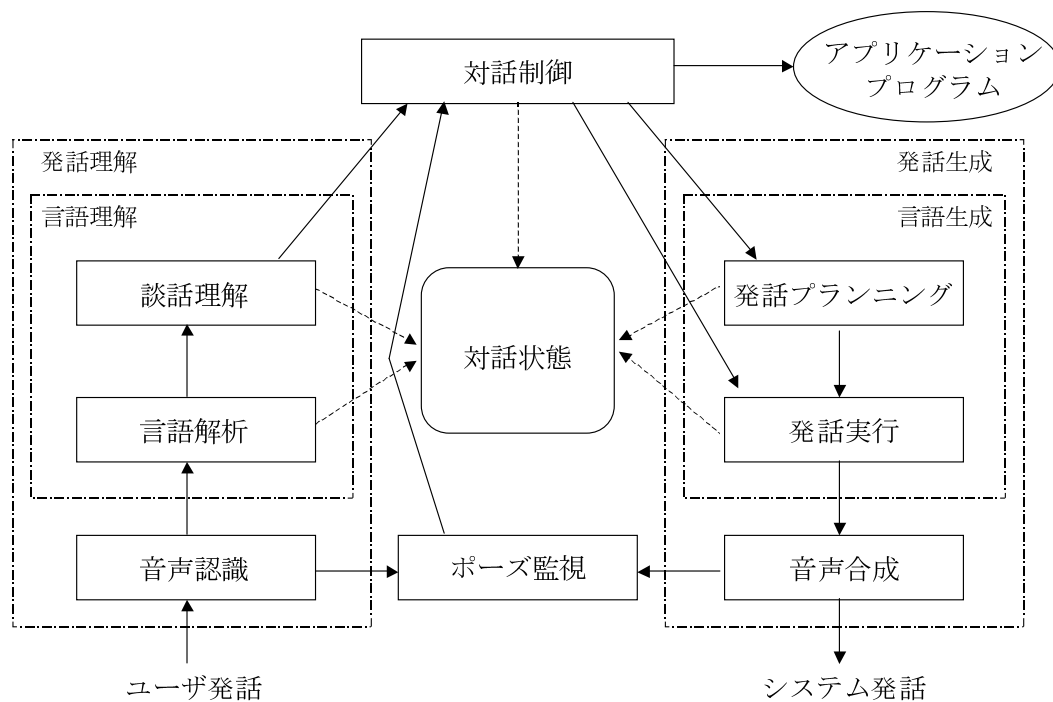


図 2.2: 音声対話システムの構成

の 2 種類のコストの和を最小にするように対話各時点のシステム行動は選択される。対話制御部が発話理解部、発話生成部と適切に連動することにより、対話の各時点において、コストの和が最小となるような対話を導くシステム行動が選択される。提案法は、従来法では回避できない不要な確認を避け、効率的な主導権混在型対話を実現できる。

2.4 本研究におけるシステム構成

本研究における音声対話システムの構成について概略を述べる。詳細は各研究テーマにおいて論じる。図 2.2 に音声対話システムのシステム構成を示す。第 1 章で述べたように、音声対話システムは、大きく分けて、発話理解部、対話制御部、発話生成部から成る。これ以外に、ポーズ¹の長さを監視するポーズ監視部と、タスク遂行のために必要となるデータベース検索などのアプリケーションプログラ

¹ 発話途中の沈黙区間をポーズと呼ぶ。

ムがある．対話制御部は必要に応じてアプリケーションプログラムを呼び出す．発話理解，対話制御，発話生成の各構成部は次の機能をもつ．

発話理解部: 入力されるユーザ音声からユーザの問い合わせ内容を理解する．

対話制御部: 対話の各時点においてシステムが達成すべき対話目標や実行すべき対話行為を決定する．

発話生成部: 対話制御部が決定する対話目標，対話行為を実現するための言語表現列を生成して，音声により出力する．

ここで，対話目標とは対話相手の信念や意図に影響を与えるという目標のことを言い，対話行為とは対話目標を達成するために成される行為のことを言う．対話行為のうち，表層の言語表現の発話に直接対応するものを表層的対話行為と呼ぶ．なお，本論文においてはシステムの対話行為を簡単にシステム行動と呼ぶ場合がある．

さらに，発話理解部は，音声認識部，言語解析部，談話理解部から成る．発話生成部は，発話プランニング部，発話実行部，音声合成部から成る．このうち，発話プランニング部は従来研究における談話生成の機能を果たす構成部であり，発話実行部は従来研究における表層生成の機能を果たす構成部である．言語解析部と談話理解部を合わせて，言語理解部と呼び，発話プランニング部と発話実行部を合わせて，言語生成部と呼ぶ．各構成部の機能を以下に説明する．

音声認識部: 入力されたユーザ音声を単語列に変換する．

言語解析部: 音声認識部が認識した単語列の語彙・統語的な特徴からユーザ発話内容を組み立てる．

談話理解部: 言語解析部が組み立てたユーザ発話内容を対話の文脈と関連付けて理解する．

発話プランニング部: 対話制御部が決定した対話目標を談話として実現するための表層的な対話行為の列 (発話プラン) を生成する．

発話実行部: 発話プランニングの決定にしたがって, 表層の言語表現を生成し, 言語表現を音声で出力するように音声合成部に対して指示する.

音声合成部: 発話実行部の指示にしたがって, 言語表現をシステム音声に変換して出力する.

本研究では, 特に, 談話理解部, 対話制御部, 発話プランニング部に焦点をあてる. 各構成部は, 対話状態と呼ばれる共有情報を参照・更新しながら, 連動して並列に動作する. 各構成部がどのように連動し, 対話状態にこういった情報が保持されているかは, 対話のフェーズに依存する.

まず, ユーザ問い合わせ把握フェーズにおける構成部の機能, 構成部間の連動, 対話状態の内容について概略を説明する. 本研究で取り上げる3つの研究テーマの中で, ユーザ問い合わせ把握フェーズに関連するのは, 語用論的条件の充足に基づく発話理解と効率的対話制御の2つの研究テーマである. 本研究では, ユーザ問い合わせ把握フェーズにおいて, 対話状態には次の情報が保持されていると考える.

- システム理解状態
- 対話登場人物間の社会関係, 話し手の視点, 情報のなわばりに関する制約
- 直前に実行されたシステムの対話行為

システム理解状態とは, ユーザ問い合わせ内容に対するシステムの理解の結果を表したものである. システム理解状態は属性と値の対の集合として表されることが多い [6]. 本研究では, 属性の値がユーザからの承認発話によって承認されたかどうかということも考慮に入れて, システム理解状態を3つ組 < 属性, 値, 承認フラグ > の集合として表す. 承認フラグは属性の値が承認済みかどうかを表す. たとえば, 天気情報案内システムの場合, 場所, 日といった属性があり, 場所属性の値としては都市の名前, 日属性の値としては今日, 明日といった値を考えることができる.

対話登場人物間の社会関係, 話し手の視点, 情報のなわばりに関する制約は, 第3章で述べる語用論的条件に充足に基づく発話理解において, ユーザ発話から抽出

した語用論的条件を充足させるための文脈として用いる．直前に実行されたシステム対話行為は，第 5 章で述べる効率的対話制御において，ユーザの承認発話によってどの情報が承認されたかを正しく理解するために必要である．

ユーザ問い合わせ把握フェーズにおいて，発話理解部は，現在の対話状態の下でユーザ発話を理解し，システム理解状態を更新する．発話理解部の中で，本研究が関心をもつのは，談話理解部の働きである．本研究では，談話理解の働きとして，ゼロ代名詞の指示対象の決定法について論じる．音声認識，言語解析という発話理解部の他の機能については議論しない．対話制御部は，第 5 章で説明するように，対話の各時点において，対話の効率性の観点から適切なシステムの対話行為を決定する．対話制御部は，発話理解部が随時更新するシステム理解状態を参照し，発話生成部に対してシステム応対フェーズにおいて予期されるシステム応答の長さを問い合わせることによって，最小の長さの対話を導くようなシステム対話行為を選択する．発話生成部は，対話制御部が選択した対話行為に相当する言語表現を生成し，音声として出力する．ユーザ問い合わせ把握フェーズにおいて，発話生成部は一連の発話を談話として生成するのではなく，確認や情報要求のための発話を単独に生成すると考えている．したがって，発話プランニングの機能は利用せず，発話実行部が対話行為に相当する言語表現を直接に生成する．

次に，システム応対フェーズにおける構成部の機能，構成部間の連動，対話状態の内容について概略を説明する．本研究で取り上げる 3 つの研究テーマの中で，システム応対フェーズに関連するのは，漸次的発話生成と効率的対話制御の 2 つの研究テーマである．このうち，効率的対話制御は，対話の効率性の観点から，ユーザ問い合わせ把握フェーズからシステム応対フェーズに移行するタイミングを決定するという機能を果たす．システム応対フェーズに移行した後は，システム応答の途中にユーザのアイツチや割り込み発話が起きることは想定せず，システム応答が終了した時点で自動的にユーザ問い合わせ把握フェーズに移行するものとして議論を進める．したがって，効率的な対話制御は，システム応対フェーズにおける構成部の連動，対話状態について特別の仕組みは必要としない．これに対して，漸次的発話生成は，システム応対フェーズにおける複雑な構成部の連動と対話状態の管理が必要となる．

システム応対フェーズにおいて漸次的発話生成を実現するための構成部の機能，構成部間の連動，対話状態の内容について概略を説明する．システム応対フェーズにおいて対話状態には以下の情報が保持されている．

- 実施中の発話プラン
- ユーザに伝達済みの情報
- ユーザ応答の履歴
- 注視状態

第4章で説明するように，アプリケーションプログラムは，ユーザ問い合わせ内容にしたがって，ユーザに伝達すべき情報を生成する．発話プランニング部は，ユーザに必要な情報を伝達するという対話目標を達成するための発話プランを生成し，対話状態に書き込む．このとき，発話プランとしては，話し言葉に特有の小さな発話単位を使って，ユーザに段階的に情報を伝達するような発話プランが立案される．発話実行部は，発話プランに基づいて言語表現を生成し，音声合成部を通してシステム音声を出力する．発話プランニングと発話実行を並行して進め，発話すべき内容が最終的に決定するのを待たずに発話を開始し，発話を実行しながら詳細な発話プランを再立案することにより，漸次的な発話生成が実現できる．個々の発話の実行が完了したら，ユーザに伝達済みの情報が対話状態に書き込まれる．注視状態 [34] はシステム発話の一部を代名詞化するためなどに用いられる．発話実行部は，注視状態の移行を伴う言語表現が発話が完了すると，注視状態追跡のためのルールに基づいて注視状態を変更する．

システム発話途中でユーザからのアイツチや割り込み発話が起きた場合には，発話理解部がその時点の対話状態にしたがってユーザ応答タイプを分類し，対話制御部に知らせる．システム対話行為のいずれに対してユーザ応答が起きたのかが対話状態に書き込まれる．対話制御部は，ユーザ応答タイプと，ユーザ応答に適切に対処するための協調的対話原則と呼ばれるルールに基づいて，必要ならば発話を中断し，発話プランを変更するように発話プランニング部に指令する．このようにして，ユーザのアイツチや割り込み発話に即座に対処し，臨機応変に話の流れを変更しながら，自然な発話を生成することが可能となる．

2.5 まとめ

本章では，本研究が焦点をあてる主導権混在型の音声対話システムについて論じ，本論文で取り上げる３つの研究テーマの位置付けを述べた．また，音声対話システムの構成について概観し，各研究テーマにおける構成部の機能，構成部間の連動，対話状態の内容について概略を述べた．

第 3 章

語用論的条件の充足に基づく発話理解

3.1 はじめに

自然言語によるコミュニケーションにおいては，話し手は聞き手に対して伝達する情報のすべてを言語表現として明示する必要はない．これは，聞き手が，発話が成された際に話し手と聞き手が共有している文脈情報を利用することによって，言語表現として明示されていない情報を補完することが期待できるためである．文脈情報に依存して伝達内容が定まる言語表現を文脈依存表現と呼ぶ．

音声対話システムにとって，文脈依存表現を適切に理解できることは，システムとユーザとの間に円滑な音声対話コミュニケーションを実現するために必要不可欠である．システムが文脈依存表現の内容を適切に理解することができなければ，ユーザは効率的でない冗長な言語表現を使わざるをえなくなり，音声対話の円滑さが阻害される．また，音声対話システムだけでなく，対話翻訳といった対話を処理しようとする試みにおいても，文脈依存表現を扱う手法の開発は重要な課題である [40]．

日本語音声対話において特徴的な現象として，伝達すべき情報の一部が頻繁にゼロ代名詞化されるという現象がある [39]．特に，対話登場人物を指示する言語表現はゼロ代名詞化されることが通常である．対話の中でのゼロ代名詞の使用例を見るために，図 3.1 に対話 (d3.1) を示す．この対話は，会議への参加申し込みに関する質問者と事務局との間の対話である．発話 (u3.1.1) では，質問者は会議に参加したいという希望を事務局側に伝えているが，「参加する」という行為の主

対話 (d3.1)

- (u3.1.1) 質問者: 会議に参加したいのですが .
- (u3.1.2) 事務局: 登録用紙に必要事項を記入する必要があります .
- (u3.1.3) 事務局: もうお持ちでしょうか?
- (u3.1.4) 質問者: 持っていません .
- (u3.1.5) 事務局: それではお送りします .

図 3.1: ゼロ代名詞の使用例

体, すなわち行為者はゼロ代名詞化されており, その指示対象は質問者である . 発話 (u3.1.2) では, 「記入する」の行為者はゼロ代名詞化されており, その指示対象は質問者である . 発話 (u3.1.3) と (u3.1.4) では, 「持つ」という事象の経験者と対象がゼロ代名詞化されており, その指示対象は, それぞれ質問者, 登録用紙である . 発話 (u3.1.5) では, 「送る」の行為者, 受益者, 対象がゼロ代名詞化されており, その指示対象は, それぞれ事務局, 質問者, 登録用紙である . この例に見るように, 日本語対話においては対話登場人物への指示はゼロ代名詞化されることが通常である . したがって, 日本語対話を扱う音声対話システムは, ユーザ発話内容を的確に把握するために, ゼロ代名詞の指示対象を的確に決定することが必要とされる .

自然言語処理, 計算言語学において, 文脈依存表現の適切な使用を説明するためのモデルの構築や, 文脈依存表現の伝達内容を理解するための方法は, 重要な課題として研究が進められてきた . 文脈依存表現の理解方法に対するアプローチには, 言語表現と文脈の関係を説明する合理的なモデルに基づくアプローチ [1, 7, 8, 34, 41, 51, 52, 86] と, 大規模なコーパスから学習される確率・統計的なモデルに基づくアプローチがある [30, 79, 81] . 本研究では, 前者の言語表現と文脈の関係を説明する合理的なモデルに基づくアプローチに関心がある .

言語表現と文脈との関係を説明する一般的な枠組として, Grosz 等が提案した言語構造, 意図構造, 注視状態の 3 つの構成要素から成る談話構造理論がある [34] .

さらに、談話構造理論を具体化した文脈依存表現理解方法として、対話参加者の意図構造をプランとして表現し、プランの認識に基づいて文脈依存表現の理解を行なう方法 [1, 8, 51, 52] や、対話参加者の注視状態のモデルの一つであるセンタリング理論に基づいて代名詞の指示対象を決定する方法 [7, 41, 86] がある。

プラン認識に基づく方法は、行為やイベント間の因果関係についての知識を前もって与えることができるなら、代名詞だけでなく、省略表現、間接的な発話行為などの様々な文脈依存の言語現象を扱うことができる強力な方法である。しかし、対話ドメインにおける行為やイベント間の因果関係についての知識は膨大な量となることが予想され、それらの知識を書き下すことには多大な労力を要するという問題点がある。

センタリング理論は注視状態の遷移を少数のルールで説明する。しかし、センタリング理論は、一つの談話セグメント内で適用することが意図されており、談話セグメントを切り出すためには、対話参加者のもつ意図構造の情報が必要となる [86]。したがって、センタリング理論に基づく方法も、ドメインにおける行為やイベント間の因果関係に関する知識を必要とするという問題から逃れることは難しい。

日本語対話においては、待遇表現 [59] や受給表現 [48]、情報のなわばり [42] に関わる様々な文末形式が頻繁に使われる。これらの言語表現が適切に使用されるためには、対話登場人物間の社会関係 [59]、話し手の視点 [48]、情報のなわばり [42] に関する特定の語用論的条件が文脈において成立していなければならない。文脈は対話登場人物間の社会関係などに関してどのような事実が成立すべきかを示す制約の集合であると考え、語用論的条件を文脈の下で充足させることにより、ゼロ代名詞の指示対象を決定できる場合がある。

たとえば、図 3.1 で示した対話 (d3.1) の発話 (u3.1.5) について考える。発話 (u3.1.5) では、「送る」の行為者と受益者はゼロ代名詞化されており、待遇表現「お送りする」が使用されている。この表現は謙譲表現であり、その使用に際しては、語用論的条件として『話し手は、「送る」の行為者より受益者を上位に待遇する』という条件が課せられる。文脈には、『話し手は話し手自身より聞き手を上位に待遇する』という社会関係に関する制約と、『話し手は事務局である』、『聞き手は質問者である』という現在の話し手、聞き手についての制約が含まれるとする。この文脈の

下で語用論的条件を充足させるためには、ゼロ代名詞化された行為者は話し手の事務局であり、受益者は聞き手の質問者と割り当てればよいことが分かる。対話登場人物間の社会関係などの語用論的条件を充足させるために用いられる制約は、ドメインの行為やイベントの因果関係に関する知識よりも、小さな労力で記述できることが期待される。本研究では語用論的条件の充足に基づいてゼロ代名詞の指示対象を決定する方法に注目する。

Yoshimoto は、待遇表現や受給表現に関する語用論的条件が日本語対話におけるゼロ代名詞の指示対象を決定することに利用できることを議論しているが、その議論は対話登場人物として話し手と聞き手の二人のみが存在する場合に限定されている [90]。また、ユーザ発話によって文脈に新規な情報が導入される場合が考慮されていない。

中岩は、日英機械翻訳システムにおいて高精度のゼロ代名詞解析を実現することを目的として、翻訳システム評価用例文の調査に基づいて、接続語、用言意味属性、様相表現の語用論的・意味論的制約に関するゼロ代名詞解析ルールを導いている [62]。この方法は、言語表現の使用に直接的に結びついている条件や制約を利用するという点で、本研究の提案法と関連がある。

本研究では、制約の集合として文脈を表現するが、これは状況理論 [4] の考え方に由来する。状況理論では、発話の情報内容と発話文脈との相互関係は、情報内容の間の制約という考え方をを用いて捉える。橋田は、状況理論において利用された制約という考え方をさらに追求し、言語にまつわる情報の部分性を扱うために、依存伝搬という制約プログラミングの方法を提示している [36]。このように、多様な発話文脈と発話内容との相互関係をモデル化するために、制約という考え方は有望である。

本章では、日本語対話において、待遇表現、受給表現、情報のなわばりに関わる特定の文末形式を使用する際に課せられる語用論的条件を文脈の下に充足させることによって、対話登場人物を指示するゼロ代名詞の指示対象を決定する発話理解法を提案する。提案法は、話し手、聞き手以外の対話登場人物が存在する場合にも適用可能であり、ユーザ発話によって新規な文脈情報が導入されることも説明できる。語用論的条件、文脈を構成する制約を論理式として表現し、論理的な推論方式に基づいて語用論的条件を文脈の下に充足させる計算の枠組みを与える。

と、ユーザ問い合わせ内容に対するシステム理解状態が保持されている．語用論的条件を充足させるための文脈は、制約の集合としてモデル化される．本研究で用いる制約としては、対話登場人物間の社会関係、話し手の視点、情報のなわばりに関する制約がある．

ユーザ発話が入力されると、ユーザ発話内容とともに語用論的条件が抽出される．本章では、語用論的条件を発話の語彙・統語論的特徴から取り出す言語解析 [46] を想定し、語用論的条件は抽出できたものとして議論を進める．また、音声認識誤りについては議論しない．

談話理解部は、語用論的条件を文脈の下で充足させることにより、ユーザ発話内容に含まれるゼロ代名詞の指示対象を決定する．このプロセスを理解プロセスと呼ぶ．理解プロセスによって情報を補完されたユーザ発話内容によって、システム理解状態は更新される．また、提案法は、発話に含まれるゼロ代名詞の指示対象を決定するだけでなく、発話によって新規に導入される文脈情報を取り出すことができる．新規に導入された文脈情報は、後続の対話におけるユーザ発話の理解やシステム発話の生成に役立てられる．

提案法において、言語表現から抽出する語用論的条件と文脈を構成する制約は状況理論における事態を使って表現される．事態は次の形をしている．

$$\langle\langle R, o_1, \dots, o_n ; p \rangle\rangle. \quad (3.1)$$

式 (3.1) で、 R は関係、 $o_1, \dots, o_n$ は対象の列、 p は極性を表す．式 (3.1) は、極性が 1 のとき、対象 $o_1, \dots, o_n$ の上に関係 R が成立することを示し、極性が 0 のとき関係 R が成立しないことを示す．関係 R の引数 o_i ($i = 1 \dots n$) には、定数かパラメタが割り当てられる．パラメタは不定項を表し、‘*’ で始まる記号として書かれる．理解プロセスの中で値が定まっていないゼロ代名詞の指示対象はパラメタとして書かれる．語用論的条件と文脈を表現するための事態を表 3.1 に示す．

事態 (3.5) は、対話参加者 x_1 から見たときに、人物 x_2 が人物 x_3 より上位に待遇されていることを示すために用いる．一方、事態 (3.6) は、対話参加者間で共通に認識されている社会関係を表すために用いる．

事態 (3.7) において、人物 x_1 が人物 x_2 に帰属するとは、 x_1 が x_2 の身近な人物であることを意味する．例えば、ある人物にとって、家族の構成員はその人物に

帰属する．こういった帰属関係は，社会関係や情報のなわばりに関する制約を定義する際に重要である．

事態 (3.8) は，話し手の視点 (共感度) [48] を表すために用いる．事態 (3.9) で使われた情報のなわばり [42] という概念については，3.3.3 節 で説明する．

理解プロセスの中で (3.10) の形の事態が現れた場合，行為 *in fon* の行為者に対応する対象が，関係 *Agent* の第一引数 x に割り当てられている対象と一致するかどうかを判定し，一致するなら，この事態が成立するとし，さもなければ，この

表 3.1: 語用論的条件と文脈を表現するための事態

$\ll Speaker, x ; 1 \gg$ (人物 x は話し手である)	(3.2)
$\ll Hearer, x ; 1 \gg$ (人物 x は聞き手である)	(3.3)
$\ll Person, x ; 1 \gg$ (対象 x は人物である)	(3.4)
$\ll Respect, x_1, x_2, x_3 ; 1 \gg$ (人物 x_1 は人物 x_2 を人物 x_3 より上位に待遇する)	(3.5)
$\ll Social_rel, x_1, x_2 ; 1 \gg$ (人物 x_1 は人物 x_2 より上位に待遇される)	(3.6)
$\ll Attrib, x_1, x_2 ; 1 \gg$ (人物 x_1 は人物 x_2 に帰属する)	(3.7)
$\ll Empathy, x_1, x_2, x_3 ; 1 \gg$ (人物 x_1 は人物 x_3 より人物 x_2 寄りの視点をとる)	(3.8)
$\ll Territory, x, in\ fon ; 1 \gg$ (情報 <i>in fon</i> は人物 x のなわばりに入る)	(3.9)
$\ll Agent, x, in\ fon ; 1 \gg$ (人物 x は行為 <i>in fon</i> の行為者である)	(3.10)
$\ll Neg, x_1, x_2 ; 1 \gg$ (x と y は異なる人物である)	(3.11)

事態は成立しないとする。また、(3.11) の形の事態が現れるときには、必ず両引数に具体的な引数が割り当てられており、両引数に割り当てられた対象が異なるかどうかを判定できると仮定する。両引数に割り当てられた対象が異なるなら、この事態は成立するとし、さもなければ、成立しないとする。

3.3 言語表現の使用と語用論的条件

3.3.1 待遇表現の使用と社会関係

日本語対話においては、待遇表現が頻繁に使用される。待遇表現としては、動詞句の尊敬型 (例:「お～になる」、「お～なさる」) や謙譲型 (例:「お～する」、「お～致す」)、接頭辞による対象物の尊敬表現 (例:「お名前」) などがある。こういった待遇表現の使用に際しては、以下に示すような語用論的条件が課せられる。

言語使用規則 1 動詞句の尊敬型の使用に際しては、話し手は主語の対象を非主語の対象より上位に待遇するという語用論的条件が課せられる。

言語使用規則 2 動詞句の謙譲型の使用に際しては、話し手は非主語の対象を主語の対象より上位に待遇するという語用論的条件が課せられる。

言語使用規則 3 接頭辞による対象物の尊敬表現においては、話し手は、その対象物の所有者を話し手自身より上位に待遇する。

例えば、図 3.1 に示した対話 (d3.1) の発話 (u3.1.5) では、謙譲表現「お送りします」が使用されているので、言語使用規則 2 により、次の語用論的条件が抽出される。ただし、「送る」のゼロ代名詞化された行為者と受益者を表すパラメタをそれぞれ $*zero_1$ と $*zero_2$ で表す。

$$\langle\langle Speaker, *sp_1 ; 1 \rangle\rangle. \quad (3.12)$$

$$\langle\langle Respect, *sp_1, *zero_2, *zero_1 ; 1 \rangle\rangle. \quad (3.13)$$

すなわち、話し手 $*sp_1$ は人物 $*zero_2$ を $*zero_1$ より上位に待遇していなければならない。

3.3.2 受給表現の使用と話し手の視点

日本語対話において受給表現の使用は、話し手の対話登場人物に対する視点(共感度)と関係があることが知られている[48]。例えば、受給表現「～てもらう」の使用に際しては、次の話し手の視点に関する語用論的条件が課せられる。

言語使用規則 4 「てもらう」の使用に際しては、話し手は間接目的語の対象より主語の対象寄りの視点をとるという語用論的条件が課せられる。

3.3.3 文末形式と情報のなわばり

情報のなわばり[42]とは、発話内で言及された事態全体を話し手がどのように眺めているかということ考察することから神尾によって考案された概念である。神尾は、ある情報が話し手のなわばりに属するか否かを決定する条件として、次のものをあげている[42]。

情報が話し手のなわばりに属する条件[42]

- (a) 情報が話し手自身の行動予定などに関するものであること。
- (b) 話し手が直接体験して得た情報であること。
- (c) 話し手及び話し手の特に身近な人物に関する情報であること。
- (d) 話し手の専門領域に関する情報であること。

さらに、神尾は、話し手のなわばりに入る情報のみが直接形によって言及されるという一般化を行なっている。

情報のなわばりについての語用論的条件が課せられる文のタイプを発語内行為タイプ[77]によって分類することを考える。発語内行為とは、言語使用の慣例にしたがって発話を行うことのうちに遂行される行為である。表 3.2 に発語内行為タイプを示す。

本研究では、情報のなわばりについての語用論的条件が課せられる文として、話し手と聞き手の間で情報の授受を行うことを目的とする文に注目し、表 3.3 に示す 5 つの文のタイプを取り上げる。タイプ 1 とタイプ 3 の文で授受される情報は、

話し手のなわばりに入り，タイプ2，タイプ4，タイプ5の文で授受される情報は，話し手のなわばりには入らないという語用論的条件が課せられる．まとめると次のようになる．

言語使用規則 5 情報の授受を目的とする次の文末形式を用いて情報に言及する際には，情報は話し手のなわばりに入らないという語用論的条件が課せられる．

- 直接的な文末タイプをもつ陳述表示型の文
- 行為拘束型の文

言語使用規則 6 情報の授受を目的とする次の文末形式を用いて情報に言及する際には，情報は話し手のなわばりに入らないという語用論的条件が課せられる．

表 3.2: 発語内行為タイプ

行為拘束型 話し手が聞き手に対して事態を達成する意図をもっていることを伝達する． 例: 「名前を申し上げます」
行為指導型 話し手が聞き手に対して事態を達成するように指導する． 例: 「お名前を教えてください」
陳述表示型 話し手が聞き手に対して事態が成立することを伝達する． 例: 「名前は堂坂浩二と申します」
疑問提示型 (a) 話し手が聞き手に対して事態が成立するか否か知りたいことを伝達する． 例: 「登録用紙はお持ちでしょうか？」 (b) 話し手が聞き手に対して特定の条件を満たす対象が何か知りたいことを伝達する． 例: 「ご住所はどちらですか」

- 間接的な文末タイプをもつ陳述表示型の文
- 疑問提示型の文
- 行為指導型の文

情報のなわばりについて例示するために、図 3.3 に対話 (d3.2) を示す。発話 (u3.2.1) では、「田中先生を招く」という事態が「～ということだそうですが」という間接形を用いて言及されているので、言語使用規則 6 より、その事態は話し手のなわ

表 3.3: 文のタイプと情報のなわばり

タイプ 1	直接的な文末タイプをもつ陳述表示型の文で情報を授受 授受情報のなわばり: 話し手 例: 「(人物 P の) 名前は堂坂浩二です」 授受情報: 「P の名前」
タイプ 2	間接的な文末タイプをもつ陳述表示型の文で情報を授受 授受情報のなわばり: 話し手以外 例: 「(人物 P の) 名前は堂坂浩二のようです」 授受情報: 「P の名前」
タイプ 3	行為拘束型の文で情報を授受 授受情報のなわばり: 話し手 例: 「(人物 P の) 名前を申し上げます」 授受情報: 「P の名前」
タイプ 4	疑問提示型の文で情報を授受 授受情報のなわばり: 話し手以外 例: 「(人物 P は) 登録用紙を持っていますか？」 授受情報: 「P が登録用紙を持っている」
タイプ 5	行為指導型の文で情報を授受 授受情報のなわばり: 話し手以外 例: 「(人物 P の) 住所を教えてください」 授受情報: 「P の住所」

対話 (d3.2)

(u3.2.1) 質問者:田中先生を会議にお招きするということだそうです。

(u3.2.2) 事務局:はい。

(u3.2.3) 事務局:基調講演でお話しになります。

図 3.3: 情報のなわばりについての対話例

ばりに入らないという語用論的条件が課せられる。また，発話 (u3.2.3) では，直接形「お話しになります」が使用されているので，言語使用規則 5 より，「基調講演でお話しになります」という事態は話し手のなわばりに入るという語用論的条件が抽出される。

3.4 文脈

文脈は制約の集合として表される。ここでは，制約の記述形式と，会議参加申し込みというドメインにおいて有効な制約を示す。

$$R_1(\overrightarrow{*x_1}) \wedge \cdots \wedge R_{n-1}(\overrightarrow{*x_{n-1}}) \Rightarrow R_n(\overrightarrow{*x_n}). \quad (3.14)$$

上の式 (3.14) で， $\overrightarrow{*x_i}$ ($i=1 \dots n$) はパラメタの列である。 $R_i(\overrightarrow{*x_i})$ ($i=1 \dots n$) は，パラメタ化された事態である。上の制約は，左辺の条件 $R_i(\overrightarrow{*x_i})$ ($i=1 \dots n-1$) がすべて充足するなら，右辺の条件 $R_n(\overrightarrow{*x_n})$ も充足することを示している。

制約 (3.14) で，左辺の条件が必要ない場合は，制約は次の形で書かれる。

$$R(\overrightarrow{*x}). \quad (3.15)$$

上の形の制約 (3.15) が，もしパラメタを含まないなら，その制約は事実と呼ばれる。事実もまた文脈を構成する制約の一種である。

会議参加申し込みに関するドメインにおいて有効な制約について説明する。まず，図 3.4 に社会関係についての制約を示す。制約 (3.16) は，対話参加者は対話参加者自身に帰属するという帰属関係の反射律を示している。制約 (3.17) は，話

$$\llcorner \textit{Attrib}, *x, *x ; 1 \gg. \quad (3.16)$$

$$\llcorner \textit{Speaker}, *sp ; 1 \gg \wedge \quad (3.17)$$

$$\llcorner \textit{Hearer}, *hr ; 1 \gg \wedge$$

$$\llcorner \textit{Person}, *less ; 1 \gg \wedge$$

$$\llcorner \textit{Person}, *more ; 1 \gg \wedge$$

$$\llcorner \textit{Neg}, *less, *more ; 1 \gg \wedge$$

$$\llcorner \textit{Attrib}, *less, *sp ; 1 \gg \wedge$$

$$\llcorner \textit{Attrib}, *more, *hr ; 1 \gg$$

$$\Rightarrow \llcorner \textit{Respect}, *sp, *more, *less ; 1 \gg.$$

$$\llcorner \textit{Speaker}, *sp ; 1 \gg \wedge \quad (3.18)$$

$$\llcorner \textit{Person}, *less ; 1 \gg \wedge$$

$$\llcorner \textit{Person}, *more ; 1 \gg \wedge$$

$$\llcorner \textit{Neg}, *less, *more ; 1 \gg \wedge$$

$$\llcorner \textit{Social_rel}, *more, *less ; 1 \gg$$

$$\Rightarrow \llcorner \textit{Respect}, *sp, *more, *less ; 1 \gg.$$

図 3.4: 社会関係についての制約

話し手は聞き手に帰属する人物を話し手に帰属する人物より上位に待遇することを示す．制約 (3.18) は，人物 **more* が人物 **less* より上位に待遇されるなら，話し手 **sp* は **more* を **less* より上位に待遇することを示している．

図 3.5 は話し手の視点と情報のなわばりについての制約を示している．まず，制約 (3.19) は，話し手は聞き手に帰属する人物よりも話し手に帰属する人物寄りの視点をとることを示している [48]．次に，制約 (3.20)，(3.21)，(3.22) は，情報のなわばりに関する制約であり，3.3.3 節で示した神尾の提案 (a) ~ (d) をこの対話ドメインに適合するように近似したものである．制約 (3.20) は，(a) の近似であり，事態 **soa* の行為者が人物 **agent* なら，その事態は，その人物の行動予定を示すものであり，その人物の情報のなわばりに入ると考えている．制約 (3.21) は，聞き手が行為者であるような事態は，話し手のなわばりには入らないことを示している．これは神尾が与えている条件 (a) をさらに強めたものになっている．制約 (3.22) は (c) に対応しており，ある人物 **agent2* のなわばりは，その人物に帰属する人物 **agent1* のなわばり内の情報を含むことを示している．なお，条件 (b) や (d) に対応する制約は与えない．そういった制約を与えるためには，ドメインの行為やイベントに関する知識が必要とされと考えられるからである．

また，各発話ごとに，話し手と聞き手が誰であるかという事態が文脈に含まれているとする．すなわち，次に示す事態のうち，(3.23) と (3.24) か，(3.25) と (3.26) かのいずれか 2 つの事態が文脈に含まれる．

《《*Speaker*, 事務局 ; 1》》 (3.23)

《《*Hearer*, 質問者 ; 1》》 (3.24)

《《*Speaker*, 質問者 ; 1》》 (3.25)

《《*Hearer*, 事務局 ; 1》》 (3.26)

さらに，次の事実が文脈に含まれているとする．

《《*Attrib*, 事務局, 質問者 ; 0》》 (3.27)

《《*Attrib*, 質問者, 事務局 ; 0》》 (3.28)

$$\ll Speaker, *sp ; 1 \gg \wedge \quad (3.19)$$

$$\ll Hearer, *hr ; 1 \gg \wedge$$

$$\ll Person, *less ; 1 \gg \wedge$$

$$\ll Person, *more ; 1 \gg \wedge$$

$$\ll Neq, *less, *more ; 1 \gg \wedge$$

$$\ll Attrib, *more, *sp ; 1 \gg \wedge$$

$$\ll Attrib, *less, *hr ; 1 \gg$$

$$\Rightarrow \ll Empathy, *sp, *more, *less ; 1 \gg.$$

$$\ll Agent, *agent, *soa ; 1 \gg \quad (3.20)$$

$$\Rightarrow \ll Territory, *agent, *soa ; 1 \gg.$$

$$\ll Speaker, *sp ; 1 \gg \wedge \quad (3.21)$$

$$\ll Hearer, *hr ; 1 \gg \wedge$$

$$\ll Agent, *hr, *in fon ; 1 \gg$$

$$\Rightarrow \ll Territory, *sp, *in fon ; 0 \gg.$$

$$\ll Attrib, *agent1, *agent2 ; 1 \gg \wedge \quad (3.22)$$

$$\ll Neq, *agent1, *agent2 ; 1 \gg \wedge$$

$$\ll Territory, *agent1, *soa ; 1 \gg$$

$$\Rightarrow \ll Territory, *agent2, *soa ; 1 \gg.$$

図 3.5: 話し手の視点と情報のなわばりについての制約

$\ll Person, \text{質問者} ; 1 \gg$ (3.29)

$\ll Person, \text{事務局} ; 1 \gg$ (3.30)

すなわち，このドメインでは，事務局と質問者が互いに他に帰属することはないことを仮定する．また，事務局を人物として扱うことも仮定する．

3.5 語用論的条件の充足に基づく発話理解プロセス

本節では，語用論的条件を文脈の下で充足させることにより発話を理解するプロセスについて説明する．

3.5.1 単一化

パラメタを含まない2つの事態が与えられるとき，それぞれの事態の関係名と極性が同一で，同じ引数位置にある対象が同一なら，2つの事態は単一化可能である．単一化した結果の事態は，元の2つの事態に含まれるすべての引数位置に対応する対象をもつ．また，極性のみが異なる2つの事態は互いに矛盾するという．

次に，パラメタを含む事態の間の単一化について説明する．パラメタへの対象の割当てをアンカーと呼ぶ．パラメタを含む2つの事態が与えられるとき，それぞれの事態に含まれるパラメタに対象を割当てることによって，2つの事態が単一化可能になるような最小のアンカーを見つける．そのようなアンカー F が存在するなら，パラメタを含む2つの事態はアンカー F の下に単一化可能であるという．

3.5.2 理解プロセス

発話から抽出された語用論的条件にはゼロ代名詞の指示対象を表すパラメタが含まれている．理解プロセスの目的は，語用論的条件を文脈の下で充足させることにより，パラメタに適切な指示対象を割り当てるアンカーを見つけることである．このためには，線形導出法 [54] を適用して，語用論的条件が制約の集合から導出できることを判定する手続きを実行すればよいと考えられる．しかし，語用

論的条件のすべてが，その時点の文脈を構成する制約から導き出せるものではなく，発話によって新規な制約が文脈に導入されることもあり得る．そういった場合を扱うためには，文脈に含まれていない制約の存在を仮定しながら解釈を進めていく必要がある．仮定された制約は，解釈が終わった後文脈に導入され，後続対話におけるユーザ発話の理解やシステム発話の生成に役立てられる．

図 3.6 に理解プロセスを示す．最初に，文から抽出された語用論的条件の列 $\vec{C}$ と文脈を構成する制約の集合 Σ が与えられる．語用論的条件の列 $\vec{C}$ に含まれるすべてのパラメタに対するアンカーを F とし，理解の結果新規に導入される新制約の集合を Δ とする．アンカー F の初期値は空であり，新制約の集合 Δ の初期値は空集合である．理解プロセスは語用論的条件の列 $\vec{C}$ が空列となるとき成功し，停止する．このときのアンカー F がゼロ代名詞の指示対象を与える．このアンカーをゼロ代名詞の理解結果と呼ぶ．

図 3.6 の (2) と (3) において， C_1 と矛盾する事実とは， C_1 と極性だけが異なるような事実である． C_1 と矛盾する事実が文脈から導き出せるかどうか判定することは，一般の線形導出法に基づいて行なう．(4) は，制約によって語用論的条件が書き換えられるプロセスであり，一般の線形導出法と同等のものである． $C_1[F]$ は，条件 C_1 にアンカー F を適用した結果得られる条件を示す．(4) で適用される制約が，左辺の条件をもたない式 (3.15) の形の場合は， $\vec{C}$ には何も付け加えない．このとき， C_1 は充足されたと言われる．(5) は， C_1 にパラメタが含まれ，適用可能な制約が存在しない場合で，このとき充足手続きは失敗する．適用可能な制約が複数ある場合には，理解プロセスを分岐させる．

上の手順で，(3) は一般の線形導出法の拡張である．(3) では，新制約を仮定することによって理解プロセスを進め，仮定された制約は理解プロセスが成功して終了したとき文脈に付け加えられる．新規に文脈に導入される制約は，パラメタを含まない事実型の制約のみである．

3.5.3 理解結果の選択

理解プロセスによって，ゼロ代名詞の指示対象は，入力発話から抽出された語用論的条件に含まれるパラメタに対象を割り当てるアンカーによって与えられる．

```

/* 変数の初期設定 */
 $\vec{C} :=$  入力発話から抽出された語用論的条件に初期設定;
 $\Sigma :=$  文脈を構成する制約の集合に初期設定;
 $F := \emptyset$ ; /*  $\vec{C}$  に含まれるパラメタに対するアンカー */
 $\Delta := \emptyset$ ; /* 理解プロセス終了後に新規に文脈に導入される新制約の集合 */
/* ループ */
While ( $\vec{C} \neq \emptyset$ ) {
(1)    $\vec{C}$  の先頭から語用論的条件  $C_1$  を取り出す
(2)   If  $C_1$  がパラメタを含まず, かつ,  $C_1$  と矛盾する事実を  $\Sigma$  から導出可能
      Then 理解プロセスを失敗させて, ループから脱出する
(3)   Else If  $C_1$  がパラメタを含まず, かつ,  $C_1$  と矛盾する事実は  $\Sigma$  から
      導出可能ではない
      Then  $C_1$  を  $\vec{C}$  から消去する
       $C_1$  を  $\Delta$  に付け加る
(4)   Else If 制約  $(R_1 \wedge \dots \wedge R_{n-1} \Rightarrow R_n)$  が  $\Sigma$  に含まれ, かつ,
       $C_1[F]$  と  $R_n$  がアンカー  $G$  の下に単一化可能
      Then  $\vec{C}$  から  $C_1$  を消去する
       $R_1, \dots, R_{n-1}$  を  $\vec{C}$  に付け加える
       $F$  に  $G$  の割当て情報を付け加える
(5)   Else 理解プロセスを失敗させて, ループから脱出する }

```

図 3.6: 理解プロセス

すべての理解プロセスが失敗したなら，その発話は理解不能であって，その発話に含まれるゼロ代名詞の指示対象は決定不能とする．さらに，成功に終る複数の理解プロセスがあるときは，次の基準に基づいて理解結果を選択する．

理解結果の選択基準

(a) できるだけ多くのパラメタに対象を割当てることができる理解結果を優先する．

(b) 理解後に文脈に導入される新制約の集合がより小さい解釈を優先する．

(b) は，発話を解釈するために新たに導入される仮説をできるだけ少なくするという方策であり，一種の儉約原則 [15] と見なすことができる．この基準によって，解釈が一意に定まるなら，図 3.6 における新制約の集合 Δ に含まれる新たな事実を文脈 Σ に付け加え，続く文の解釈に進む．また，この基準によっても理解結果を一意に定められない場合には，文内のゼロ代名詞の指示対象は不定となる．

3.5.4 ゼロ代名詞の指示対象の決定例

提案法によって日本語ゼロ代名詞の指示対象がどのように決定されるかを例示する．例として図 3.3 に示した対話 (d3.2) の発話 (u3.2.1) と (u3.2.3) を用いる．対話 (d3.2) を会議参加申込みを受け付ける音声対話システムとユーザの対話と見なす場合には，質問者の発話がユーザ発話，事務局の発話がシステム発話となり，システムが理解すべきはユーザ発話のみということになるが，ここでは，提案法の能力を見るために，質問者，事務局の発話の双方を理解する過程を示す．

まず，発話 (u3.2.1) において，ゼロ代名詞化された「招く」の行為者の指示対象について考える．このゼロ代名詞の適切な指示対象は事務局である．この発話が成されるとき，文脈には話し手と聞き手が誰かを示す事実 (3.25) と (3.26) が含まれている．発話からは次の語用論的条件 (3.31) ~ (3.34) が抽出される．なお，発話 (u3.2.1) の「招く」のゼロ代名詞化された行為者をパラメタ $*zero_3$ で表す．

$$\langle\langle Speaker, *sp_2 ; 1 \rangle\rangle \quad (3.31)$$

$$\langle\langle Person, 田中 ; 1 \rangle\rangle \quad (3.32)$$

$$\ll \textit{Respect}, *sp_2, \text{田中}, *zero_3 ; 1 \gg \quad (3.33)$$

$$\ll \textit{Territory}, *sp_2, \ll \text{招く}, *zero_3, \text{田中}, \text{会議} ; 1 \gg ; 0 \gg \quad (3.34)$$

語用論的条件 (以下, 条件と略す)(3.33) は, 言語使用規則 2 より, 謙讓表現「お招きする」の使用から抽出される．条件 (3.34) は, 言語使用規則 6 より, 間接形「～ということだそうです」の使用から抽出される．

まず, 文脈に含まれる事実 (3.25) によって, 条件 (3.31) は充足し, パラメタ $*sp_2$ の値は質問者となる．また, 条件 (3.32) と矛盾する事実は文脈から導出できないので, 条件 (3.32) は新制約の集合に付け加えられる．

さらに, パラメタ $*sp_2$ の値が決まったことから, 条件 (3.33) は次の (3.35) に書き換えられている．

$$\ll \textit{Respect}, \text{質問者}, \text{田中}, *zero_3 ; 1 \gg \quad (3.35)$$

条件 (3.35) には制約 (3.17) と (3.18) が適用できる．まず, 制約 (3.17) が適用されるとすると, 結局, 次の条件が導き出される．

$$\ll \textit{Person}, *zero_3 ; 1 \gg \quad (3.36)$$

$$\ll \textit{Person}, \text{田中} ; 1 \gg \quad (3.37)$$

$$\ll \textit{Neq}, *zero_3, \text{田中} ; 1 \gg \quad (3.38)$$

$$\ll \textit{Attrib}, *zero_3, \text{質問者} ; 1 \gg \quad (3.39)$$

$$\ll \textit{Attrib}, \text{田中}, \text{事務局} ; 1 \gg \quad (3.40)$$

このうち, 上の条件 (3.36) は, 事実 (3.29) か (3.30) と単一化可能である．まず, 事実 (3.29) と単一化されたとすると, パラメタ $*zero_3$ は質問者となる．このとき, 条件 (3.38) は充足され, 条件 (3.39) は次のように書き換えられている．

$$\ll \textit{Attrib}, \text{質問者}, \text{質問者} ; 1 \gg \quad (3.41)$$

条件 (3.41) は制約 (3.16) によって充足する .

一方 , 条件 (3.36) が事実 (3.30) と単一化された場合について考える . このとき , パラメタ $*zero_3$ の値は事務局となり , 条件 (3.39) は , 次のように書き換えられている .

$$\ll \textit{Attrib}, \textit{事務局}, \textit{質問者} ; 1 \gg \quad (3.42)$$

条件 (3.42) は , 事実 (3.27) と矛盾する . したがって , 条件 (3.36) が事実 (3.30) と単一化された場合には , 理解プロセスは失敗することが分かる .

結局 , 条件 (3.35) に制約 (3.17) を適用した結果 , パラメタ $*zero_3$ の値は質問者でなければならないことが分かる .

次に , 条件 (3.34) の充足を考える . 今 , パラメタ $*zero_3$ が質問者と同定されているので , 条件 (3.34) は次のように書き換えられている .

$$\ll \textit{Territory}, \textit{質問者}, \ll \textit{招く}, \textit{質問者}, \textit{田中}, \textit{会議} ; 1 \gg ; 0 \gg \quad (3.43)$$

しかし , この条件 (3.43) と矛盾する事実が , 制約 (3.20) を用いて導けるので , 条件 (3.35) に制約 (3.17) を適用した場合には . 理解プロセスは失敗することが分かる . しかし , 条件 (3.35) には制約 (3.18) も適用できる .

条件 (3.35) に制約 (3.18) を適用すると , 次の条件が得られる .

$$\ll \textit{Person}, *zero_3 ; 1 \gg \quad (3.44)$$

$$\ll \textit{Person}, \textit{田中} ; 1 \gg \quad (3.45)$$

$$\ll \textit{Neq}, *zero_3, \textit{田中} ; 1 \gg \quad (3.46)$$

$$\ll \textit{Social_rel}, \textit{田中}, *zero_3 ; 1 \gg \quad (3.47)$$

条件 (3.44) は事実 (3.29) か (3.30) と単一化可能である . まず , 事実 (3.29) と単一化されたとすると , パラメタ $*zero_3$ は質問者となる . しかし , パラメタ $*zero_3$ を質問者と同定して理解プロセスを進めても , 既に議論したように , 条件 (3.34) が文脈と矛盾してしまう . よって , 条件 (3.44) が事実 (3.30) と単一化した場合のみを考えればよい . このとき , パラメタ $*zero_3$ は事務局と同定され , 条件 (3.46) は充足され , 条件 (3.47) は次のように書き換えられている .

《《*Social_rel*, 田中, 事務局 ; 1》》 (3.48)

条件 (3.48) と矛盾する事実は文脈から導けないので，条件 (3.48) は新制約の集合に付け加えられる．このとき，条件 (3.34) は，次のように書き換えられている．

《《*Territory*, 質問者, 《招く, 事務局, 田中, 会議 ; 1》 ; 0》》 (3.49)

条件 (3.49) と矛盾する事実は文脈から導き出せないので，条件 (3.49) は新制約の集合に付け加えられる．なお，条件 (3.49) は，制約 (3.21) を使えば充足させることができるが，提案法では，パラメタを含まない条件で，文脈に矛盾しない条件は新制約としてみなされるので，条件 (3.49) は新制約とみなされる．

結局，発話 (u3.2.1) から抽出された語用論的条件を文脈の下で充足させることによって，パラメタ $*zero_3$ の値が適切に事務局と同定された．さらに，新たな事実 (3.32)，(3.48)，(3.49) が文脈に導入される．

次に，発話 (u3.2.3) のゼロ代名詞化された「お話しになる」の行為者の指示対象を決定することを考える．このゼロ代名詞の適切な指示対象は，発話 (u3.2.1) で導入された人物「田中」である．発話 (u3.2.3) が成されるとき，文脈には事実 (3.23) と (3.24) が含まれている．発話 (u3.2.3) からは次の条件が抽出される．ただし，「お話しになる」のゼロ化された行為者をパラメタ $*zero_4$ で表す．

《《*Speaker*, $*sp_3$; 1》》 (3.50)

《《*Respect*, $*sp_3$, $*zero_4$, $*sp_3$; 1》》 (3.51)

《《*Territory*, $*sp_3$, 《話す, $*zero_4$, 基調講演 ; 1》 ; 1》》 (3.52)

上の条件 (3.51) は尊敬表現「お話しする」の使用から抽出される．また，条件 (3.52) は，言語使用規則 5 より，「 $*zero_4$ が基調講演で話す」という事態が「お話しになる」という直接形によって言及されていることから抽出される．

ここで，条件 (3.50) は事実 (3.23) によって充足され，パラメタ $*sp_3$ の値は事務局となる．このとき，条件 (3.51) は，次のように書き換えられている．

《《*Respect*, 事務局, $*zero_4$, 事務局 ; 1》》. (3.53)

条件 (3.53) には制約 (3.17) と (3.18) が適用できる．今，制約 (3.17) を適用したとする．この制約適用によって，パラメタ $*zero_4$ の値は聞き手の質問者に帰属する人物であることがわかり，さらに，新制約を導入しない限り，それは質問者であると決定される．このとき，条件 (3.52) は次のように書き換えられている．

$$\ll Territory, 事務局, \ll 話す, 質問者, 基調講演 ; 1 \gg ; 1 \gg \quad (3.54)$$

しかし，条件 (3.54) と矛盾する事実が制約 (3.21) を用いて導出されてしまう．そこで，条件 (3.53) に制約 (3.18) を適用する．発話 (u3.2.1) で導入された新事実 (3.48) を用いることにより，パラメタ $*zero_4$ の値は田中となる．これはゼロ代名詞の適切な指示対象を与える．このとき，条件 (3.52) は次のように書き換えられている．

$$\ll Territory, 事務局, \ll 話す, 田中, 基調講演 ; 1 \gg ; 1 \gg \quad (3.55)$$

条件 (3.55) はパラメタを含まず，かつ，条件 (3.55) と矛盾する事実は導けないので，新制約の集合に付け加えられる．

結局，発話 (u3.2.3) から抽出された語用論的条件を文脈の下で充足させることにより，「お話しする」の行為者は「田中」とであると適切に同定された．また，3.3.3 節の (d) で述べたように，神尾 [42] によれば，話し手の専門領域に属する情報は話し手のなわばりに入る．誰かが基調講演で話すという事態は事務局の専門領域に属することであるから，文脈に導入される新たな事実 (3.55) は妥当なものであるといえる．

3.6 評価

提案法の有効性を実証するために，ATR 自動翻訳電話研究所で収集された「国際会議への参加申し込みに関する問い合わせ」についての対話 [39] を用いて，提案法の評価を行なった．

評価に用いた対話は，キーボード対話 [39] によって収集された 10 個の対話である．評価は，対話に含まれる各文から抽出した語用論的条件に提案法を適用することによって行った．語用論的条件の抽出は，表層の語彙・統語論的特徴から語用論的条件を取り出す解析器 [46] によって行った．

提案法を会議の問合せを受け付ける音声対話システムへの適用することを考えるなら，質問者の発話のみを理解することに注目すればよいが，ここでは，手法の適用対象とは独立に手法の能力を検証するために，提案法を質問者，事務局の発話の双方に適用することを試みた．

表 3.4 に (a) 評価用対話データの内訳，(b) 提案法による結果，(c) ゼロ代名詞の指示対象として話し手を強制的に割り当てた場合の結果を示す．(a) の評価用対話データの内訳は，対話データに含まれる対話文の個数，ゼロ代名詞の個数，対話登場人物を指示するゼロ代名詞の個数を示している．ここでは，表層でゼロ化されている構成要素のうち，発話内容を把握するために補完される必要があるもの（主語，目的語，間接目的語など）をゼロ代名詞と認定した．対話データに現れたゼロ代名詞のうち，25 個（157 個- 132 個）のゼロ代名詞は，対話登場人物以外の対象を指示するもので，これらは提案法の適用範囲外である．また，質問者と事務局以外の第三者を指示するゼロ代名詞は存在しなかった．

(b) の提案法による結果は，対話登場人物を指示するゼロ代名詞のうち，提案法

表 3.4: 語用論的条件の充足に基づく発話理解法の評価結果

(a) 評価用対話データの内訳

対話文数	ゼロ代名詞数	人物指示の ゼロ代名詞数
225	157	132

(b) 提案法による結果

人物指示の ゼロ代名詞数	正解数	誤り数	対象外数
132	108	1	23

(c) 指示対象を強制的に話し手に割り当てた場合の結果

人物指示の ゼロ代名詞数	正解数	誤り数	対象外数
132	71	61	0

対話 (d3.3)

(u3.3.1) 質問者:ホテルの手配もして頂けるのですか?

(u3.3.2) 事務局:京都ホテルは予約できます.

対話 (d3.4)

(u3.4.1) 質問者:会議の間に市内観光があるそうですが.

(u3.4.2) 事務局:はい, 8月5日の午後に清水寺, 金閣寺などを見学します.

図 3.7: 提案法が指示対象を決定できなかったゼロ代名詞の例

が正しく指示対象を決定できたゼロ代名詞の数 (正解数), 提案法が誤った指示対象を選択したゼロ代名詞の数 (誤り数), 提案法がいかなる指示対象も割り当てることがなかったゼロ代名詞の数 (対象外数) を示している.

(c) は, 対話登場人物を指示するゼロ代名詞について, 指示対象を強制的に話し手に割り当てた場合の正解数, 誤り数, 対象外数を示している. この場合, 対象外数は 0 である. 質問者と事務局以外の第三者を指示するゼロ代名詞が存在しなかったため, ゼロ代名詞の指示対象を強制的に聞き手に割り当てた場合の正解数と誤り数は, (c) の場合の逆となる.

(b) に示す提案法による結果から分かるように, 提案法は, 対話登場人物を指示する 132 個のゼロ代名詞のうち, 108 個 (82%) のゼロ代名詞について正しく指示対象を同定できた. 提案法が誤った指示対象を選択してしまったのは 1 例のみであった. この他の 23 個のゼロ代名詞については, 提案法はいかなる指示対象も割り当ててことはなかった. (c) のように, ゼロ代名詞の指示対象を強制的に話し手に割り当てるとすると, 71 個 (54%) のゼロ代名詞については正しく指示対象を同定できるが, 61 個 (46%) のゼロ代名詞については誤った指示対象が選択された. 提案法による結果 (b) と, 指示対象を強制的に話し手に割り当てた場合 (c) を比較すると, 提案法が, 対話登場人物を指示するゼロ代名詞のうち, 82% のゼロ代名詞の指示対象を正しく決定できたことは良好な結果であると言える.

提案法がいかなる指示対象も割り当てることができなかつたり, 誤った指示対

象を割り当ててしまったのは、待遇表現などの特定の言語表現が使用されていない場合や、ドメインのイベントや行為に関する知識がなければ扱うことができないゼロ代名詞が現れた場合である。それぞれの典型的な対話例を図 3.7 に示す。

図 3.7 において、対話 (d3.3) は、提案法がいかなる指示対象も割り当てることがなかった場合である。発話 (u3.3.2) の「予約できます」の行為者がゼロ化されている。しかし、この文には、待遇表現などの特定の表現が用いられていないため、提案法は「予約できます」の行為者を同定することはできない。「予約できます」という表現のかわりに「ご予約することができます」といった待遇表現が使われていたとしたら、ゼロ代名詞を適切に処理できる。

次に、対話 (d3.4) は提案法が誤った指示対象を選んでしまった例である。発話 (u3.4.2) では「見学します」の行為者がゼロ化されている。しかし、このゼロ代名詞の指示対象は、対話の中に既に導入されている対話参加者ではなく、発話 (u3.4.1) で導入された「市内観光」というイベントから想起される「市内観光の参加者」という主体であると考えられる。このゼロ代名詞を扱うためには、そういったドメインのイベントに関する知識が必要となる。こういった知識を用いない提案法は、このゼロ代名詞の指示対象について誤った予測をしてしまう。すなわち、提案法では、「見学します」という直接形で言及された事態は話し手のなわばりに入るとしており、且つ、行為者が話し手であるような行為は話し手のなわばりに入っているので、「見学する」の行為者は事務局であると誤って同定してしまう。

提案法は、ドメインに特有のイベントや行為に関する知識を用いないにもかかわらず、対話登場人物を指示するゼロ代名詞について、82%のゼロ代名詞の指示対象を正しく決定できた。対話登場人物以外の対象を指示するゼロ代名詞や、ドメインのイベントや行為に関する知識を用いなければ処理できないゼロ代名詞を扱うためには、プラン認識やセンタリング理論に基づく他の手法が必要となる。本研究の主張は、提案法がそういった既存の手法に置き換わるものであるということではなく、3.1 節で述べたように、従来法が前提としていたドメインのイベントや行為の因果関係に関する膨大な知識に頼らずとも、待遇表現などの表層表現の使用に際して課せられる語用論的条件を利用することによって、対話登場人物を指示するゼロ代名詞の多くの部分を扱うことができるということである。

3.7 まとめ

日本語対話において，待遇表現，受給表現，特定の文末形式の使用に際して課せられる語用論的条件を文脈の下に充足させることにより，対話登場人物を指示するゼロ代名詞の指示対象を決定する発話理解法を示した．ATR 自動翻訳電話研究所で収集されたキーボード対話において，提案法は 82 % のゼロ代名詞を正しく解析できた．提案法は以下の利点をもつ．

- ドメインの行為やイベントの因果関係に関する膨大な量の知識に頼ることなく，ドメインにおいて有効な対話登場人物間の社会関係，話し手の視点，情報のなわばりに関する制約を与えることにより，対話登場人物を指示するゼロ代名詞の指示対象を決定できる．
- ユーザ発話に含まれるゼロ代名詞の指示対象を文脈情報に基づいて決定するだけでなく，ユーザ発話によって導入される新規な文脈情報を取り出すことができる．
- 発話理解部と発話生成部が文脈情報を共有して連動することにより，ユーザ発話やシステム発話によって導入される文脈情報を，後続の対話におけるユーザ発話の理解だけでなく，適切なシステム発話生成に役立てることが可能となる．

今後の課題としては以下のことがある．

- 提案法は，対話登場人物以外を指示するゼロ代名詞や，待遇表現などの特定の言語表現が使われていない発話には適用できない．そういった場合に対処するためには，提案法で利用した対話登場人物間の社会関係などの文脈情報と，従来より利用されてきた意図構造，注視状態といった他の文脈情報を統合した発話理解法を構築する必要がある．提案法では論理的な制約の集合として文脈を表現した．同様の論理的な制約に基づく枠組みの中で，対話参加者の意図構造，イベント間の時間関係といった他の文脈情報も表現できることが示されているが [18, 20]，文脈情報の中には，代名詞と指示対象候補の間の文間距離のように，選好に基づいてモデル化することが相応しい文脈情

報もある．また，複数の文脈情報を利用すると，異なる文脈情報が互いに矛盾する理解結果を導く場合が生じる．このように，多様な文脈情報を統合した発話理解法を構築するためには，論理的な制約だけでなく，選好に基づいて文脈情報を表現する枠組みや，複数文脈間の矛盾に対処する方法が必要となる．

- 提案法では，文の理解が曖昧なときには，3.5.3 節で述べた戦略にしたがって理解を一意に絞込む．しかし，対話においては，ある発話の理解の曖昧さが後続の発話の理解結果との整合性に基づいて解消される場合がある．こういった場合を扱うためには，ある発話の解釈が曖昧であっても，それを一意に定めなくて，理解プロセスを文間を越えて分岐させ，後続発話の理解結果によって先行する文の理解を絞り込めるように，提案法を拡張することが必要である．音声対話システムにおいて，対話の各時点の理解結果の曖昧さを保持し，対話の流れの中で曖昧さを解消する方法論が提案されてきており [37, 65]，今後の課題として，そういった方法論の中に提案法を組み込んでいくことが考えられる．

第 4 章

漸次的発話生成

4.1 はじめに

音声対話は、複数の対話参加者が互いに協調しながら従事する実時間の協同活動である [12]。このため、音声対話システムは、発話を生成するとき、ユーザに対して一方的に情報を伝達するだけでは不十分であり、対話の中でやり取りされる情報についてユーザとの間で共通理解を維持しながら、発話を生成する必要がある。音声対話において、共通理解を維持しつつ対話を進めることは、対話参加者が従事すべき最も基本的な協同作業である [13]。

主導権混在型の音声対話システムがシステム対応フェーズにおいて発話を生成するとき、ユーザはシステムが伝達する情報に対してアイツチや割り込み発話を行う場合がある。システムは、発話途中のユーザのアイツチや割り込み発話に対処することによって、伝達情報についてユーザとの間で共通理解を維持しつつ発話を生成しなければならない。このとき、システムは、発話する前にいくらでも時間をかけて発話内容について熟考して良いわけではなく、発話すべきときに即座に発話を開始することが望ましい。以上の観点から、本章では、主導権混在型の音声対話システムが、システム対応フェーズにおいて、システム発話途中のユーザのアイツチや割り込み発話に即座に対処し、ユーザ意図にしたがって話の進め方を変更しながら、自然な発話を生成するという問題に着目する。こういった発話生成によって、システム対応フェーズにおける主導権の柔軟な交代が可能となり、システムとユーザの間に円滑な対話を実現することができる。

本研究では、発話生成の中でも談話生成に焦点をあてる。談話生成とは、発話内容を決定し、発話内容を伝達するための談話構造を決定するプロセスである。自然言語生成の分野において、これまで多くの談話生成手法が提案されてきた [38, 55, 57, 58, 60, 67]。これらの従来法は、書き言葉における文を単位として談話を生成することに注目している。しかし、話し言葉による音声対話では、書き言葉における文よりも小さな発話単位が頻出することが分かっている [16, 22, 35, 49, 68, 70]。このことは、書き言葉の文よりも小さな単位が音声対話コミュニケーションの最小単位であり、音声対話システムとユーザとの対話においても、文よりも小さな発話単位に対してユーザのアイツチや割り込み発話が生じる可能性があることを意味している。文を単位として発話を生成すると、ユーザのアイツチや割り込み発話に対して即座に対処できず、文のどの部分に対してユーザのアイツチや割り込み発話が成されたのか適切に判断できない。したがって、書き言葉の文を単位として発話を生成する従来法は、システム発話途中のユーザからのアイツチや割り込み発話に即座に対処しながら、自然な発話を生成することは難しい。

ユーザからのアイツチや割り込み発話に即座に対処しながら自然な発話を生成するための一つの方策は、話し言葉に特有の小さな発話単位を使って、発話を漸次的に生成することである。漸次的な発話生成とは、「発話すべき内容を考えながら同時に発話を生成し、発話を行いながら次に発話すべき内容を決定していくこと」を言う。小さな発話単位を使って、ユーザの反応を観察しながら、段階的に情報を伝達していくことにより、システム発話途中のユーザのアイツチや割り込み発話に対して即座に対処することが可能となる。加えて、発話内容を組み立てながら発話を生成することにより、発話内容を最終的に決定する前に発話を開始できるので、発話すべきときに即座に発話を開始することが容易となる。実際、人は日常の音声会話において発話を漸次的に生成するものと考えられている [49, 68]。人にせよ、システムにせよ、あらゆることを一時に推論できるわけではなく [85]、推論能力に限界のある話し手が、発話内容を完全に具体化してから発話を開始できるわけではないのは当然である。

本章では、タスク指向型対話において、システムがユーザのアイツチや割り込みで即座に対処し、ユーザ意図にしたがって臨機応変に話の進め方を変更しながら、自然な発話を生成することを可能とする漸次的発話生成法を示す。そのため

に、まず、人同士の対話を書き起こした対話データにおいて、発話単位と談話構造を分析し、音声対話では話し言葉特有の小さな発話単位に基づく微細な談話構造が頻出し、そういった微細な談話構造が段階的な情報伝達に役立っていることを示す。次に、発話途中に起きる対話相手からの応答へ対処するための談話戦略を分析する。これらの分析から、システムがユーザからの応答に対処しながら発話を生成するために必要となる協調的対話原則を導く。

本研究で提案する漸次的発話生成法は、話し言葉特有の小さな発話単位を使って、一連の発話を一つの談話として生成する。小さな発話単位を使うことにより、ユーザに段階的に情報を伝達できるので、ユーザからのアイツチや割り込み発話に即座に対処できる。発話内容の計算と発話の実行は並行して進められ、発話内容が完全に決定するのを待たずに、発話が可能となった時点で話し言葉特有の小さな発話単位を使って発話を開始する。このことにより、システムが発話すべきときに、即座に発話を開始することが容易となる。協調的対話原則にしたがって、発話生成部が発話理解部、対話制御部と適切に連動することにより、発話途中のユーザのアイツチや割り込み発話に即座に対処しながら発話を生成し、ユーザの意図にしたがって、臨機応変に話の進め方を変更することが可能となる。

漸次的発話生成の計算モデルとしては、漸次的な文形成に関心を絞ったモデルが従来より提案されてきている [17, 45]。これらのモデルの主な関心は、発話内容が与えられた後に文の統語構造の形成や言語表現の合成を文頭から文末に向かって漸次的に進めることにある。漸次的な統語構造形成や言語表現合成は重要な問題であるが、本研究では立ち入らない。本研究の関心は、発話内容の決定を段階的に進めながら、話し言葉特有の小さな発話単位を使って、一連の発話を自然な談話として漸次的に生成することにある。

POPEL は、漸次的発話生成の実働システムである [73]。POPEL は主として POPEL-WHAT と POPEL-HOW という 2 つの構成部から構成される。POPEL-WHAT は発話内容を決定し、段階的に POPEL-HOW に送る。POPEL-HOW は段階的に文を形成する。POPEL は文脈情報を参照しながら文の列を一つの談話として生成することができる [74]。しかし、POPEL の目的は主として書き言葉の生成であり [74]、話し言葉対話に特有の微細な談話構造は考慮されていない。本研究では、話し言葉対話に特有の微細な談話構造に基づいて漸次的に談話を生成す

る方法を提案する．

実時間対話において言い淀みや言い直しが起きることを説明するために，対話の実時間性を考慮した発話生成システムの枠組が提案されている [9]．実時間における発話生成を考慮するという点に関して，本研究の提案法はこの従来の枠組みと共通点がある．しかし，この従来の枠組みでは，実時間の中で発話を生成するシステムの構想が示されているだけで，具体的な計算モデルは示されていない．本研究では，階層的プランニング手法に基づいて，漸次的発話生成の具体的な計算モデルを示す．

ユーザ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 [10, 60]．これらの従来システムは，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行った場合に，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という対処法を備えている．本研究においても，システム発話生成途中に起きるユーザからの応答に対処する際に，これらの従来システムと同等の方法を用いる．本研究の特徴は，話し言葉対話特有の小さな発話単位でのやり取りの中で，ユーザからの応答に対処しながら発話を生成することにある．

従来の主導権混在型の音声対話システムの中には，システム発話へのユーザ割り込み (バージン) を許す従来システム [43] が存在する．しかし，従来システムは，ユーザ割り込みに対してシステム発話を中断するのみであり，ユーザ割り込みに応じてシステムの話の流れを適切に変更することはできない．また，システム発話へのユーザ割り込みを許したとしても，書き言葉における文を単位として発話を出力している限り，ユーザ割り込みがシステム発話のどの部分に対して成されたのかを正しく判断することができず，ユーザ割り込みに対して適切に対処することができない．本研究における提案法では，話し言葉特有の小さな発話単位を使って段階的に情報を伝達することにより，ユーザ割り込みに対して即座に対処することができ，ユーザ割り込みがシステム発話のどの部分に対して成されたのかを適切に判断することが可能となる．

本章では，まず，人同士の会話を書き起こした対話データにおいて，話し言葉特有の発話単位と談話構造，発話途中に起きる対話相手からの応答へ対処するための談話戦略について分析し，分析結果に基づいて，システムがユーザ応答に対処

対話 (d4.1)

- (u4.1.1) 説明者: 愛甲石田まで来てもらって/<はい>
(u4.1.2) でそこでバスなんですけど/<はい>
(u4.1.3) えーっとですね バスーの行き先が/
(u4.1.4) えっとー 森の里青山行きという<はい>バスがあるので/<はい>
(u4.1.5) それに乗ってもらって/<はい>

図 4.1: 書き起こされた対話からの抜粋

しながら発話を生成するために必要となる協調的対話原則を導き出す。次に、システムが、協調的対話原則にしたがって、ユーザのアイツチや割り込み発話に対処しながら、漸次的に発話を生成することを可能とする漸次的発話生成法を示し、提案法の有効性について考察する。

4.2 対話データに基づく協調的対話原則の分析

発話単位、談話構造、対話相手からの応答へ対処するための談話戦略という観点から、人同士のタスク指向型対話を書き起こした対話データを分析した。実験には 90 人の被験者が参加した。各対話において、二人の対話参加者 (質問者と説明者) は電話で会話することにより「質問者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が要請された。さらに、対話の現実味を高めるために、質問者は対話終了後目的地に実際に行くことが要請された。質問者と説明者は、説明者が問題を解くために十分な知識をもっており、質問者はそのような知識をもっていないように選ばれた。80 対話が収集され書き起こされた。すべての対話に関して、質問者が対話の中で獲得した交通経路で目的地に到着できたことが確認されている。

対話の書き起こしは文献 [78] で提案された記法にしたがって行なわれた。図 4.1 は書き起こされた対話の一部を示している。この対話では、説明者が交通経路を

説明するために発話している．説明者の発話の途中で現れる<はい>という記法は，その時点で質問者が「はい」という間投詞的表現を発話したことを示している．記号“/”は次の4.2.1節で説明する発話単位の区切りを示している．収集した対話の中から分析のために20対話を選んだ．

4.2.1 発話単位の分析

話し手は，伝達すべき情報全体を一塊として発話を生成するのではなく，情報を複数の単位に分配して，その単位ごとに発話を生成するものと考えられる [35, 49, 80]．ここではそのような単位を発話単位と呼ぶ¹．発話単位は談話構造の最小単位とみなすことができる．

発話単位の切り出しには表層の言語表現だけでなく情報の新旧の区別や音調といった手掛かりも使う必要があるが，ここでは第一近似として次の条件を使って発話単位を切り出した．

- 節は発話単位である．
- 間投的表現は発話単位である．
- つなぎ語は発話単位の区切りである．
- 言い直しは発話単位の区切りである．

対話の書き起しにおいては，記号“/”によって発話単位の区切りを示した．図4.1に示した対話では，説明者による談話は5個の発話単位から構成されている²．

発話単位の認定においては，間投用法の終助詞 [56] で終る句も節とみなした．たとえば，次の発話において「そこからですね」は節である．

例：「そこからですね/えーとバスに乗って」

¹こういった単位の呼称は文献によって異なる．たとえば，処理単位 [49]，情報単位 [35, 80] といった呼称がある．

²発話単位 (u4.1.4) に見られるような「名詞 + という」という言語形式は定型的な表現と考えて引用節とはみなさなかった．

間投的表現とは「はい」といったアイツチなどに使われる表現で，つなぎ語とは「えーと」，「あの一」といった語である．ある話し手の発話単位を区切るときには，その話し手自身が発した間投的表現のみを考慮した．言い直しとしては次の場合を考えた．

- 同じ語や句が繰り返される場合

例：「バス/バスの乗り場がありまして」

- 発話が途中で中断され，直後に訂正される場合

例：「通信研究所/N T Tの通信研究所ってのが」

分析した対話データには 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 [22]．名詞句や後置詞句を発話した直後につなぎ語が発話されると，そこで一つの発話単位が終るので，名詞句または後置詞句の 1 つ以上の並びが発話単位となりえる．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 個の名詞単位が現れた．

話し言葉生成システムを設計する際，発話単位の大きさを決定することは重要なことである．そこで対話データに現れた名詞単位と節単位に関して，1 単位中に現れる名詞句の数を数えた．従属名詞句をもつ名詞句は全体で 1 個と数えた．節単位と名詞単位に関して，1 単位中に現れた名詞句の数の頻度分布を表 4.1 に示す．名詞単位と節単位に関して，1 単位中に現れた名詞句の数は，平均が 1.01 であり，分散は 0.26 であった．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．たとえば，図 4.1 に示した発話 (u4.1.3) は「えっとー」というつなぎ語が後続するので一つの発話単位と見なせる．書き言葉では発話 (u4.1.3) と (u4.1.4) を併せて一つの節として実現することが普通であると考えられる．

表 4.1: 名詞単位と節単位に関して，1 単位中に現れる名詞句の数の頻度分布

名詞句数	0	1	2	3	4	合計
頻度	179	1221	166	13	1	1580

情報の新旧の区別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 [35] .
そこで、情報の新旧という観点から発話単位の大きさについて分析した。まず、発話単位における情報の新旧の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したのは、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」), 対話の中で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、新情報を担う名詞句を 2 個以上含む単位は 60 個 (4%) だけであった。したがって、次の原則を導く。

協調的対話原則 1 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。

この原則を遵守して発話を生成するためには、話し言葉特有の小さな発話単位ごとに発話を生成する必要がある。

4.2.2 談話関係の分析

話し手は発話単位ごとに発話を生成する際、各発話を関係付けて一連の発話を一つの談話として生成する。この発話間の関係を談話関係と呼ぶ。談話構造とは談話関係によって互いに関係付けられた発話の階層構造である。

対話データに現れる談話構造を分析した。談話関係としては、修辞構造理論 [55] で提供される関係を用いた。ここでは文献 [38] に習って談話関係を意味関係と意図関係に分類した³。意味関係とは発話の情報内容の間の関係を分類したものである。意図関係とは、話し手がその関係を使うことによって聞き手の心的状態にどのような影響を与えることを意図しているかということ进行分类したものである。図 4.2 は図 4.1 で示した談話に現れた談話関係を図示したものである。4.2.1 節で述べた話し言葉特有の小さな発話単位が談話構造の最小単位となって、微細な談話構造がつくられていることが分かる。

³本論文における意味関係と意図関係は文献 [38] における semantic relation と interpersonal relation にそれぞれ対応する

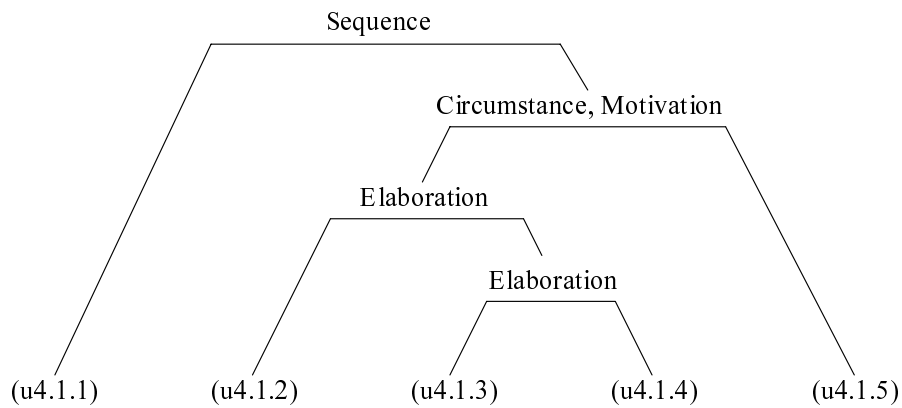


図 4.2: 図 4.1 に示した談話における談話関係

表 4.2 は談話関係の頻度分布を示す．頻繁に現れた談話関係として，Sequence，Elaboration，Circumstance，Motivation が漸次的発話生成に果たす役割を論じる．まず，Sequence 関係は主としてドメインの行為を時間順に記述していくために使われている．図 4.1 に示した対話 (d4.1) では，話し手 E は愛甲石田まで移動するという行為を記述した後，愛甲石田からバスで移動するという行為を記述している．ドメインプランを構成するすべての行為の内容が確定していなくても，Sequence 関係によってドメインプランの最初の方の部分の説明を開始できるという意味で，Sequence 関係は漸次的発話生成に貢献すると言える．しかし，Sequence 関係は話

表 4.2: 談話関係の頻度分布

(a) 意味関係

Elaboration	300
Sequence	74
Circumstance	60
Condition	26
Result	26
Purpose	2
Contrast	1

(b) 意図関係

Motivation	66
Background	14
Evidence	10
Interpretation	6
Concession	3
Enablement	1

し言葉対話に特有の小さな発話単位の間での談話関係ではない。以下に述べる他の3つの談話関係は話し言葉対話に特有の微細な談話構造をつくり、漸次的発話生成に貢献する。

Elaboration 関係は、ドメインの事象や対象の内容の一部を記述する発話と、詳細な内容を記述する発話の間での関係である。典型例は対話 (d4.1) の発話単位 (u4.1.2) と発話単位列 (u4.1.3), (u4.1.4) の間での関係である。この関係を使うことによって、話し手は発話対象の内容を完全に決定する前に発話を開始できる。

Circumstance 関係は、ドメインの事象や行為を記述する発話と、その事象や行為の環境を記述する発話の間での関係である。特に、ある行為の前提条件が成立していることを発話した後で、行為を対話相手に提案するという場合が 41 例あった。これは Circumstance 関係が使われた場合の 68%にあたる。行為の前提条件は行為の環境の一つと考えることができる。この典型例は発話単位列 (u4.1.2), (u4.1.3), (u4.1.4) と発話単位 (u4.1.5) の間での関係である。発話単位 (u4.1.5) で「バスに乗る」という行為を提案する前に、発話単位列 (u4.1.2), (u4.1.3), (u4.1.4) で「そのバスが存在する」という行為の前提条件が記述されている。行為の前提条件を記述する際には行為の構成要素に言及することが多い。話し手はこのことを利用して、行為の内容を2つ以上の発話単位に分配し、行為の内容を段階的に伝達することができる。

Motivation 関係は、行為を提案する発話とその行為の採用を聞き手に動機付けるための材料となる事実を記述する発話との間での関係である。特に、行為の前提条件が成立していることをまず言った後に行為の提案を行う場合が 41 例あった。これは Motivation 関係が使われた場合の 62%にあたる。この典型例は発話単位列 (u4.1.2), (u4.1.3), (u4.1.4) と発話単位 (u4.1.5) の間での関係である。この場合には、Motivation 関係と Circumstance 関係が同時に成立する。Motivation 関係を使うことは、Circumstance 関係を使う場合と同様に、漸次的発話生成に貢献する。

4.2.3 対話相手の応答に対処するための談話戦略の分析

音声対話において、対話の主導権 [44, 87, 88] を保持して、発話を生成している対話参加者が対話相手からの応答に対処する際に用いる談話戦略という観点から

対話データを分析した．分析結果に基づいて，システムがユーザの応答に対処するための談話戦略の指針となる協調的対話原則を導く．最初に，対話相手から応答がない場合に対処するための談話戦略について分析し，続いて，対話相手から応答があった場合に対処するための談話戦略について述べる．

以下の分析では，対話の主導権をもっている対話参加者の発話のみを対象とした．ただし，対話の主導権をもつ対話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ていない．また，発話直後に修復された発話は対象としない．発話が修復されたということは，少くとも話し手は何らかの理由でその発話を無効にして新たな発話に置き換えたと見なせる．話し手が発話直後に別の発話に置き換えた発話には，何らかの不都合があった可能性があるので，協調的対話原則を導くための分析の対象とすべきではないと考えた⁴．

まず，対話の主導権をもっている対話参加者が発話単位を生成するとき，発話単位の統語カテゴリと情報の新旧が，発話単位への対話相手からの応答の有無に与える影響を χ^2 検定を使って分析した．発話単位への対話相手からの応答として

表 4.3: 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

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

⁴ここでは，4.2.1 節で示した 1144 個の節単位と 436 個の名詞単位のすべてを分析対象としたのではなく，ここで述べた条件を満たす発話に現れる発話単位のみを分析対象とした．

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。分析結果を表 4.3 に示す。

テーブル χ^2 値が棄却域内にあるということは、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と密接な関係にあることを意味している。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。たとえば、旧情報を担う名詞単位に対して応答があった場合は、観測度数 21、期待度数 73.6、 χ^2 値 37.6 である。十分に大きな 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起しやすく、観測度数が期待度数よりも小さいならその組合せは生起しにくいことが示されている。

この分析結果から、旧情報を担う単位には対話相手からの応答がないことが期待され、新情報を担う節単位には対話相手からの応答があることが期待されることが分かる。したがって、次の協調的対話原則を導く。

協調的対話原則 2 対話相手が旧情報を担う発話単位に応答しない場合には、計画通りに発話を継続せよ。

協調的対話原則 3 対話相手が新情報を担う節単位に応答しない場合には、次に何を発話すべきかについて再考せよ。

次に、協調的対話原則 3 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 からの応答の有無と、対話参加者 P_1 が後続談話を展開する方法の間の関係进行分析した。 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、対話参加者 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 が対話の主導権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。

この分析においては、談話の展開の方法を修辞構造理論 [55] における核、衛星という概念に基づいて分類した。まず、核、衛星という概念について説明する。修辞構造理論においては、2 つの発話が談話関係によって関連付けられるとき、主たる情報を伝達する発話を核と呼び、補足的な情報を伝達する発話を衛星と呼ぶ。たとえば、図 4.1 で示した対話において、発話単位列 (u4.1.2)、(u4.1.3)、(u4.1.4) と発話単位 (u4.1.5) との間には談話関係として Circumstance 関係と Motivation 関係が成立するが、いずれの談話関係においても発話単位 (u4.1.5) が核であり、発

話単位列 (u4.1.2), (u4.1.3), (u4.1.4) が衛星となる．核，衛星という概念に基づいて，発話単位 U に後続する談話の展開の種類を次のように分類した．

拡大: U の核を記述するための後続談話が始まる．

精錬化: U の衛星を記述する後続談話が始まる．

話題移行: その他の談話展開．

拡大とは，発話単位 U で先に補足的な情報を伝達し，後続談話で主たる情報を伝達する場合である．たとえば，図 4.1 の対話において，発話単位 (u4.1.4) の後続談話の展開の種類は拡大展開と分類される．精錬化とは，発話単位 U で先に主たる情報を伝達し，後続談話でそれを補足する情報を伝達する場合である．4.2.2 節で論じたように，話し言葉対話では，修辞構造理論における Elaboration 関係を使って，抽象的な情報を核として記述した後に，詳細な情報を衛星として記述するという談話展開が頻繁に見られるが，これは精錬化展開の一種である．たとえば，発話単位 (u4.1.2) の後続談話の展開の種類は，この種の精錬化展開である．話題移行とは拡大展開でも精錬化展開でもない談話展開である．修辞構造理論における Sequence 関係のように，核と衛星の区別のない関係で発話単位と後続談話が関係付けられる場合は話題移行展開と分類される．この典型例は，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．たとえば，図 4.1 の対話において，発話単位 (u4.1.1) の後続談話の展開の種類は話題移行と分類される．

新情報を担う発話単位 U のへの応答の有無とその後の談話展開との関係を分析した結果を表 4.4 に示す．ここでは，対話参加者が新情報を担う発話単位を生成した後，対話相手からの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，まず，新情報を担う発話単位に対して対話相手からの応答がない場合には，話題移行展開が起こりにくいことが分かる．これに対して，対話相手からの応答がない場合に拡大または精錬化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，発話単位 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対話相手から承認されるまで話題移行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」) や承認のための発話 (例: 「わかりました」) を行った場合か，(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，次の協調的対話原則を導く．

協調的対話原則 4 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，節単位 U の内容を詳細化した発話単位か，節単位 U の核または衛星を記述する発話単位を生成し続けよ．

協調的対話原則 4 は，協調的対話原則 3 とともに，対話相手からの応答がない場合でも，システムは談話戦略を変更する必要がある場合を示している．この原則は，ユーザの沈黙がユーザの不理解または未承認を含意する場合があることを示している．音声対話においては，対話相手からの応答だけでなく，対話相手から応答がないことに対しても適切に対処することにより，対話相手との間で共通理解を維持する必要がある．

次に，対話の主導権をもつ対話参加者が発話している際中对話相手からの応答があった場面に注目し，対話相手から応答に対処するための談話戦略について調べた．対話相手からの典型的な応答として，承認のための肯定的な間投的表現 (例: 「はい」「ええ」)，発話単位に含まれる名詞の復唱がある．復唱には，発話単位に

表 4.4: 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

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

含まれる名詞を完全に復唱する場合 (完全復唱) と部分的に復唱 (部分復唱) する場合がある。また、名詞 *Noun* を含む発話に対する応答として成された「*Noun*+判定詞+疑問の終助詞」の形式の発話は復唱とみなした。たとえば、「愛甲石田ですか」といった発話がこれにあたる。

まず、肯定的な間投的表現に対する談話戦略について述べる。ある対話参加者 P_1 が発話単位 U を発話した後に、対話相手 P_2 が発話 U に対して肯定的な間投的表現で応答した場合を調べた。20 対話の中でそのような場合が 631 例あった。その中で、対話相手 P_2 が肯定的な間投的表現で応答した後、それに対して対話参加者 P_1 が肯定的な間投的表現を返した場合が 30 例 (5%) あった。他の場合は、対話参加者 P_1 は、対話相手 P_2 の肯定的な間投的表現による応答に対して特別の応答を行うことなく、発話を継続した。したがって、次の協調的対話原則を導く。

協調的対話原則 5 対話相手が承認のための肯定的な間投的表現を発話した場合には、発話を計画通りに継続せよ。

次に復唱に対処するための談話戦略について述べる。対話参加者 P_1 の発話単位 U に対して対話相手 P_2 が復唱 R を行った場合を考える。復唱 R が発話単位 U に含まれる名詞 *Noun* の完全復唱であった場合が 82 例あり、復唱 R が発話単位 U に含まれる名詞 *Noun* の部分復唱であった場合が 28 例あった。完全復唱と部分復唱を合わせて、対話相手 P_2 が復唱が行った場合は 110 例あった。

対話相手 P_2 の復唱 R に対して対話参加者 P_1 が採った談話戦略には次の 3 つの戦略があった。

承認: 承認のための肯定的な間投的表現を発話する。

繰り返し: 以下のいずれかを行う。

- (a) 名詞 *Noun* を繰り返す。
- (b) 名詞 *Noun* の部分 S を発話する。ただし、 S は R と S の接続が *Noun* となるような表現である。
- (c) 名詞 *Noun* によって記述された対象の内容を詳細化する。

黙諾: 復唱 R に対処するための発話は陽には行わず、そのまま発話を継続する。

それぞれの内訳は，承認が 65 例，繰り返しが 25 例，黙諾が 20 例であった．すなわち，対話参加者 P_1 が対話相手 P_2 の復唱 R に対して陽に発話を行うこと（承認あるいは繰り返し戦略）によって対応する場合が，全体の 82% を占めた．

この分析から，対話の主導権をもっている対話参加者 P_1 が，対話相手 P_2 に情報を伝達するために発話を行っているとき，対話相手 P_2 が対話参加者 P_1 の発話に対して復唱を行い，対話参加者 P_1 はその復唱に対して承認あるいは復唱によって応ずるという対話のパターンが頻出することが分かる．この対話のパターンが頻出するということは，話し手は一方的に情報を伝達するのではなく，対話相手の復唱に対処することによって，伝達情報について共通理解を維持しながら対話を進めていることを意味する．したがって，次の協調的対話原則を導く．

協調的対話原則 6 対話相手が，発話単位 U への応答として， U に含まれる名詞の復唱を発話する場合には，対話相手の復唱に対して承認，および，繰り返し戦略をとれ．

システムが書き言葉の文を単位として発話を生成すると想定すると，書き言葉の文は話し言葉特有の発話単位よりも大きいので，文の途中に発生する対話相手からの応答に対処して談話戦略を臨機応援に変更することは難しい．したがって，ここで導いた協調的対話原則を遵守しながら発話を生成するためには，システムは話し言葉特有の小さな発話単位で漸次的に発話を生成する必要がある．

4.3 漸次的発話生成法の概観

図 4.3 に漸次的発話生成のシステム構成を示す．システムは，問題解決，発話プランニング，発話実行，音声合成，対話制御，音声認識，言語理解の各構成部から成る．

ここでは，タスク指向型対話において，解くべき問題が与えられ，システムが問題の解をユーザに提案するために発話を行うという状況を想定している．提案法のシステム構成における問題解決部の役割は，ユーザに対する伝達内容全体を組み立てることである．問題解決部の実装はタスクによって異なる．たとえば，天気情報案内タスクの場合は，問題解決部はデータベースへの検索プログラムとし

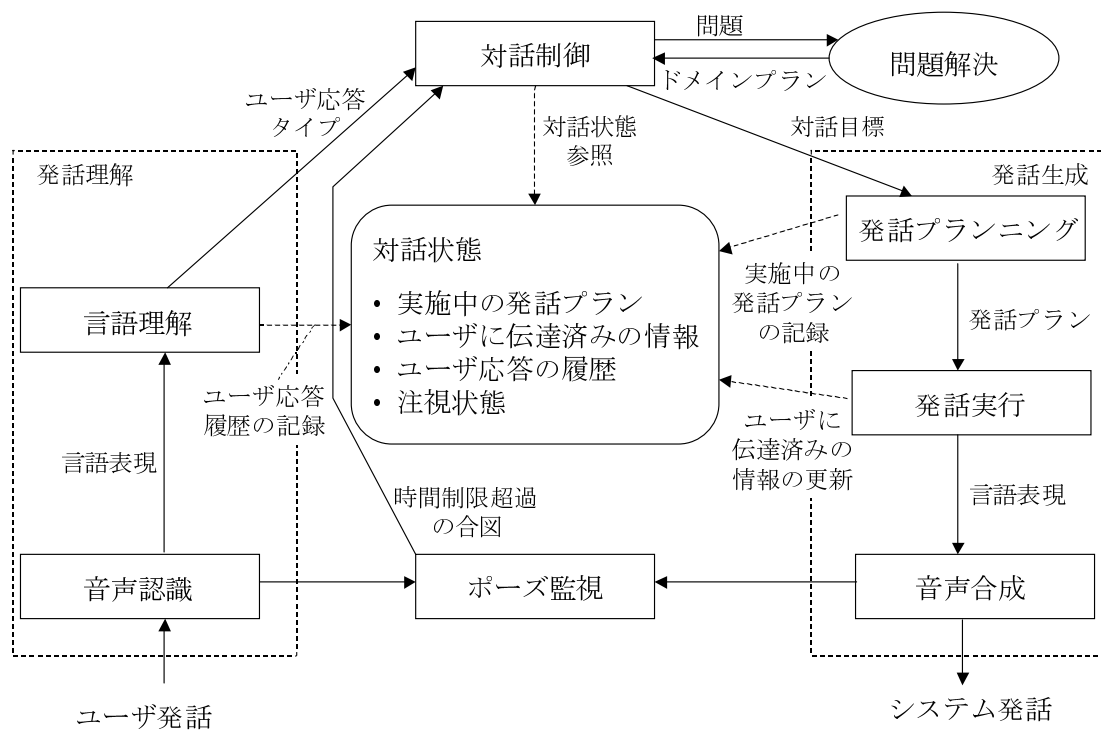


図 4.3: 漸次的発話生成のシステム構成

て実装されるであろう。本章では、タスクとして交通経路案内を取り上げ、問題解決部は、問題(交通経路探索)が与えられると、問題の解としてドメインプラン(交通経路)を立案するものとして実装され、ドメインプランは、階層的プランニングの技法 [76] を用いて、段階的に詳細化されることを想定する。

発話プランニングは、談話生成を担当する構成部であり、ドメインプランの内容をユーザに提案するという対話目標を達成するための発話プランを生成する。発話プランは、対話相手に情報を伝達することを目的とする対話行為の列である。発話プランは、階層的プランニング技法 [76] を用いて、段階的に詳細化される。最終的には、話し言葉に特有の小さな発話単位に対応する表層的な対話行為の列にまで詳細化され、発話実行部に送られる。

発話実行部は発話プラン内の表層的な対話行為に対応する言語表現を生成し、音声合成部に送る。本章では、発話単位の表層言語生成はテンプレートマッチング法 [72] によって実現するものとする。音声合成部は音声として言語表現を出力する。音声合成部は、それぞれの発話単位の発話が完了したら、そのことを発話実行部に知らせる。ポーズ監視部は、ポーズが一定の時間制限を越えたら、そのことを対話制御部に知らせる。

各構成部は、対話状態と呼ばれる情報を共有し、対話状態の参照・更新を行う。対話状態には、実施中の発話プラン、ユーザに伝達済みの情報、ユーザ応答の履歴、注視状態 [34] が書き込まれる。生成される談話が適切であることと、ユーザからの応答に適切に対処することを保証するために協調的談話原則が用いられる。協調的対話原則としては、これまでに説明した対話データの分析に基づく原則の他に、4.4 節で述べる言語学的な知見に基づく原則も用いる。

各構成部は並列に動作する。既に述べたように、ドメインプランと発話プランは階層的プランニング技法 [76] に基づいて段階的に詳細化される。ドメインプランが完全に具体化されるのを待たずに、部分的に決定されたドメインプランの内容をユーザに伝達するための発話プランが立案される。発話プランは、表層的な対話行為の列にまで段階的に具体化され、表層的な対話行為が得られた時点で発話実行部に送られ、発話が開始される。発話が行なわれている間にドメインプランは詳細化されていく。ドメインプランが詳細化された時点で発話プランは再立案される。古い発話プランに基づく発話は中断され、新しい発話プランに基づく

発話が再開される．このように，発話内容が完全に決定するのを待たずに，話し言葉特有の小さな発話単位による発話が可能となった時点で発話を開始できるので，システムは発話すべきときに即座に発話を開始することが容易となる．以上の方法によっても時間制限内に発話を開始することができず，ポーズ監視部によってポーズが一定の時間制限を越えたことが検出された場合には，対話制御部は発話生成部につなぎ語（例：「えーっと」）を出力するように命じる⁵．

システム発話の途中にユーザからのアイツチや割り込み発話が発生すると，発話生成部は，発話理解部，対話制御部と連動して，協調的対話原則にしたがってユーザのアイツチ，割り込み発話に対処する．発話理解部は，現在の対話状態に依存してユーザ応答タイプを分類し，対話制御部に知らせる．システムのどの対話行為に対してユーザ応答が起きたのかが対話状態に書き込まれる．対話制御部は，応答タイプと協調的対話原則にしたがって，必要なら発話を中断し，発話プランを変更するように発話プランニング部に指令する．システムは話し言葉特有の小さな発話単位を使って段階的に情報を伝達しているので，ユーザのアイツチや割り込みで即座に対処でき，協調的対話原則にしたがって臨機応変に話の進め方を変更しながら，自然な発話を生成することができる．

4.4 協調的対話原則と対話状態

提案方法では，これまでに述べた対話データに基づく協調的対話原則に加えて，言語学的な知見に基づく次の協調的対話原則も用いる．

協調的対話原則 7 伝達済みの情報は繰り返さない．

協調的対話原則 8 注視対象を代名詞化する．

協調的対話原則 9 注視状態に対して適切な発話を行う．

⁵問題解決部がドメインプランを段階的に詳細化できない場合であっても，発話プランニング部は，発話プランを段階的に詳細化し，発話プランが完全に決定される前に，話し言葉特有の小さな発話単位で発話を開始できるので，本システム構成によって，システムは発話すべきときに即座に発話を開始することが容易となる．

これらは Grice による協調の原則 [33] の一例とみなすことができる．協調的対話原則 7 と協調的対話原則 8 は，整然と話せ，簡潔な言い方をせよという Grice の様態の原則の一例であり，協調的対話原則 9 は，現在の対話目的と関連のあることを話せという関連性の原則の一例である．

対話状態には，発話プラン，ユーザに伝達済みの情報，ユーザ応答の履歴，注視状態 [34] が書き込まれる．発話プランニング部は発話プランを対話状態に書き込む．発話実行部は，発話プランに含まれる個々の対話行為の実行が完了したら，その対話行為が実行済みであることを対話状態に書き込む．このことによってユーザに伝達済みの情報が何かを判断することができる．発話理解部は，システムの発話に対してユーザからの応答が発生したら，ユーザ応答がどのシステム対話行為に対して行われたかを対話状態に書き込む．

注視状態の追跡は，注視状態の移行を伴う言語表現ごとに，注視状態を追跡するためのルールを書くことによって行っている．たとえば，交通経路案内の対話において場所 S から路線 L で場所 D まで移動するという行為の内容を「行く」，「出る」といった主動詞を使って記述するための一連の発話を行うとする．この一連の発話の実行された後では，注視される対象は D に移行する．また，この一連の発話が注視状態に対して適切であるためには，発話を開始した時点で場所 S が注視されている必要がある．もし場所 S が注視されていないなら，場所 S を提題化する必要がある．

4.5 問題解決

問題解決とは与えられた問題を解くドメインプランを立案することである．ドメインプランはドメイン行為の列として表される．交通経路案内に関する対話では，ドメイン行為は主としてある場所から別の場所まで交通機関を使って行くという移動行為である．ドメインプランは階層的プランニングの技法を使って段階的に詳細化されるとする．問題解決のメカニズムは漸次的発話生成手法の理解に直接関与しないので，その詳細には立ち入らない．ここでは，図 4.4 に示す路線図において武蔵野センタから厚木センタまで行くという問題が与えられた場合を例にとって，問題解決部の振舞いを例示する．


```

manner(a3, x5), dest(a3, x6)}).
cont(x4, {type(x4, station), nearest(x4, x6)}).
cont(x5, {type(x5, bus)}).
cont(x6, {type(x6, building), named(x6, “厚木センタ”)}).

```

次に，問題解決部はより詳細なドメインプランの立案を進める．与えられた問題を解く二つ以上のドメインプランが可能なら，最短時間で目的地点まで移動できるドメインプランを選択する．図 4.4 に示す路線図においては，選択されるドメインプランは，武蔵野センタからバスで吉祥寺へ移動し，そこから井の頭線の下北沢へ移動し，そこから小田急線で愛甲石田へ移動し，そこからバスで厚木センタへ移動するという 4 つの行為 a_4, a_5, a_6, a_7 から成るプランである．このプランは次の (4.3) で表され，行為 a_4 の情報内容は (4.4) で表される．他の 3 つの行為 a_5, a_6, a_7 の内容を記述するための表現は簡単のため省略する．

$$\text{plan}([a_4, a_5, a_6, a_7]). \quad (4.3)$$

$$\begin{aligned} &\text{cont}(a_4, \{\text{type}(a_4, \text{move}), \text{source}(a_4, x_1), \\ &\quad \text{manner}(a_4, x_2), \text{dest}(a_4, x_7)\}). \\ &\text{cont}(x_7, \{\text{type}(x_7, \text{station}), \text{named}(x_7, \text{“吉祥寺”})\}). \end{aligned} \quad (4.4)$$

4.6 発話プランニング

発話プランニングとは対話目標を達成する発話プランを立案することである．発話プランは対話行為の列として表される．発話プランニング部に与えられる初期の対話目標は，問題解決部が立案したドメインプランを対話相手に採用させるという対話目標である．発話プランは階層的プランニングの技法を用いて段階的に詳細化され，最終的には，言語表現の発話に対応する表層的な対話行為の列にまで詳細化される．

4.6.1 対話目標

システムが対話の流れの中で適切な発話を生成するためには、何のために発話を行なうのかという対話目標を記録しておく必要がある [60]。本研究の提案法で用いる対話目標とその表記法を以下に示す。ただし、本章を通じて対話参加者を表す引数は簡単のため省略する。

- $persuaded_plan(P)$: ユーザにドメインプラン P を採用させる。
- $persuaded_act(A)$: ユーザにドメイン行為 A を採用させる。
- $described_event(E, C, At)$: ドメイン事象 E を記述する。ただし、 E が内容 C をもつことと、話し手が E に対して態度 At をもつことが記述されている必要がある。
- $described_obj(O, C)$: ドメイン対象 O を記述する。ただし、 O が内容 C をもつことが記述されている必要がある。
- $described_thematic_rel(R, O, E)$: ドメイン対象 O がドメイン事象 E に対して格関係 R をもつことを記述する。

例えば、ドメインプラン (4.1) が立案されたとき、発話プランニング部に与えられる初期の対話目標は以下の通りである。

$$persuaded_plan([a1, a2, a3]). \quad (4.5)$$

4.6.2 対話行為

対話行為とは対話相手に情報を伝達することを目的として実行される行為である。対話行為としては、言語表現の発話に対応する表層的な対話行為と、複数の表層的な対話行為を談話関係の下に関連づけて実行することに対応する表層的でない対話行為がある。表層的な対話行為としては以下を用いる。

- $surface_desc_event(E, C, At)$: ドメイン事象 E を記述するために言語表現を発話する。ただし、その言語表現は、 E が情報内容 C をもつことと、システムが E に対して態度 At をもつこととを伝達できなければならない。

- $surface_desc_obj(O, C, R, T)$: ドメイン対象 O を記述するための言語表現を発話する。ただし、その言語表現は、 O が情報内容 C をもつことと、 O が型 T のドメイン事象に対して格関係 R をもつこととを伝達できなければならない。

表層的でない対話行為は、次の 4.6.3 節で説明する行為スキーマによって定義される。

4.6.3 階層的プランニング技法による発話プランニングの実現法

階層的プランニング技法 [76] に基づく発話プランニングの実現法について述べる。まず、発話プランニングの際に用いる知識の種類について説明する。一般に、階層的プランニングは行為スキーマと分解メソッドを使って行われる⁶。行為スキーマは、対話目標を達成する対話行為を定義するために使い、分解メソッドは、表層的でない対話行為をより具体的な対話行為の列に分解するために用いる。

行為スキーマは行為記述、適用条件、効果から成り、次のように書かれる⁷。なお、適用条件は必要がないなら省略される。

$$\begin{aligned} & \text{Act(行為記述,} \\ & \quad \text{Appcon: 適用条件,} \\ & \quad \text{Effect: 効果).} \end{aligned} \tag{4.6}$$

ここで、行為記述は変数を含む項として書かれる。適用条件と効果は論理式として書かれる。効果は対話行為によって達成される対話目標を表す。行為記述と適合する対話行為を適用条件が真となる状況で実行すると、その実行後には効果が真となる。

分解メソッドは行為記述、適用条件、プランから成り、次のように書かれる。

$$\begin{aligned} & \text{Decomp(行為記述,} \\ & \quad \text{Appcon: 適用条件,} \\ & \quad \text{Plan: プラン).} \end{aligned} \tag{4.7}$$

⁶文献 [76] では、行為スキーマはオペレータスキーマと呼ばれる。

⁷本論文では、簡単のため、行為スキーマを書く際に行為の前提条件は考慮しない。

分解メソッドの記述において、行為記述と適用条件は行為スキーマと同様である。プランは行為記述のリストであり、リストに書かれた順に行為を実行するというプランとして解釈される。適用条件が真となる状況で行為記述と適合する対話行為を実行するためには、分解メソッドに定義されたプランを実施すれば良い。

発話プランニングは、まず、行為スキーマの定義を用いて、与えられた対話目標を効果として達成する対話行為を見つけることから始まる。次に、分解メソッドを使って、その対話行為をさらに具体的な対話行為列に分解していき、最終的には、対話目標を達成する表層的な対話行為の列を得る。対話行為の列は、行為の実行順に並んでいる。分解メソッドによる対話行為の分解は、左再帰深さ優先方式で行い、次に実行する表層的な対話行為の生成の立案を優先する。

次に、話し言葉に特有な談話構造を利用した発話プランニングについて説明する。対話相手にドメインプランを採用させるという対話目標が与えられたとき、発話プランは次の方略にしたがって立案されていく。

- (1) ドメインプランを構成する個々のドメイン行為を順に対話相手に採用させていくという発話プランを立案する。
- (2) ドメイン行為の個々の構成要素を格関係とともに記述していくという発話プランを立案する。
- (3) ドメイン行為の構成要素が存在するという命題を陳述してから、その構成要素を格関係とともに記述するという発話プランを立案する。
- (4) ドメイン行為を提案した後で、そのドメイン行為の採用を動機付ける命題を陳述する。

(1) は Sequence 関係に基づく発話プランニングであり、(2) は Elaboration 関係に基づく発話プランニングである。(3) は Circumstance 関係に基づく発話プランニングであり、ドメイン行為の個々の構成要素を記述する際に用いられる。(4) は Motivate 関係に基づく発話プランニングであり、特に、対話相手からの応答に対処するために用いられる。

以下において、(1)、(2)、(3)、(4)の発話プランニングの実現法を順に示す。

$$\begin{aligned}
& \text{Act}(\text{propose_acts_in_seq}(*P), \\
& \quad \text{Appcon:plan}(*P), \\
& \quad \text{Effect:persuaded_plan}(*P)).
\end{aligned}
\tag{4.8}$$

$$\begin{aligned}
& \text{Decomp}(\text{propose_acts_in_seq}([*A \mid *R]), \\
& \quad \text{Appcon:*R} \neq [], \\
& \quad \text{Plan:[achieve(persuaded_act}(*A)), \\
& \quad \quad \text{propose_acts_in_seq}(*R)]).
\end{aligned}
\tag{4.9}$$

図 4.5: Sequence 関係に基づく発話プランニングのための諸定義

Sequence 関係

図 4.6 は, Sequence 関係に基づいてドメインプランを対話相手に採用させるための行為スキーマと分解メソッドを示す. “*” で始まる記号は変数を表す. $\text{achieve}(P)$ は目標 P を達成する行為を表す. 記法 $[H \mid R]$ は, 第一要素が H で第一要素を除く残部のリストが R であるようなリストを表す.

行為スキーマ (4.8) は Sequence 関係に基づいてドメインプランを対話相手に採用させることを目標とする対話行為を定義しており, 分解メソッド (4.9) は (4.8) の対話行為を実行するための詳細なプランを定義している.

例として, 初期の対話目標 (4.5) に (4.8) と (4.9) を適用すると, 次の発話プラン (4.10) が立案される.

$$\begin{aligned}
& \text{achieve}(\text{persuaded_act}(a1)), \\
& \text{achieve}(\text{persuaded_act}(a2)), \\
& \text{achieve}(\text{persuaded_act}(a3)).
\end{aligned}
\tag{4.10}$$

このように, (4.8) と (4.9) の適用によって, ドメインプランを構成する各行為を順にユーザに採用させていくという Sequence 関係に基づく発話プランを立案できる.

$$\text{Act}(\text{propose_act}(*A), \text{Effect: } \text{persuaded_act}(*A)). \quad (4.11)$$

$$\begin{aligned} \text{Decomp}(\text{propose_act}(*A), \text{Appcon: } \text{cont}(*A, *C), \\ \text{Plan: } [\text{achieve}(\text{described_event}(*A, *C, \text{proposal}))]). \end{aligned} \quad (4.12)$$

$$\begin{aligned} \text{Act}(\text{describe_event_by_elaboration}(*E, *C, *At), \\ \text{Effect: } \text{described_event}(*E, *C, *At)). \end{aligned} \quad (4.13)$$

$$\begin{aligned} \text{Decomp}(\text{describe_event_by_elaboration}(*E, *C, *At), \\ \text{Appcon: } *ThematicRel \in *C \wedge \\ *ThematicRel = .. [*R, *E, *O] \wedge *R \neq type \wedge \\ \text{cont}(*O, *ObjC) \wedge *Rest = *C - \{ *ThematicRel \}, \\ \text{Plan: } [\text{achieve}(\text{described_obj}(*O, *ObjC)), \\ \text{achieve}(\text{described_thematic_rel}(*R, *O, *E)), \\ \text{achieve}(\text{described_event}(*E, *Rest, *At))]). \end{aligned} \quad (4.14)$$

$$\begin{aligned} \text{Act}(\text{describe_obj_with_thematic_rel}(*O, *C, *R, *E), \\ \text{Effect: } \text{described_obj}(*O, *C) \wedge \\ \text{described_thematic_rel}(*R, *O, *E)). \end{aligned} \quad (4.15)$$

$$\begin{aligned} \text{Decomp}(\text{describe_obj_with_thematic_rel}(*O, *C, *R, *E), \\ \text{Appcon: } \text{cont}(*E, *EventC) \wedge \text{type}(*E, *T) \in *EventC, \\ \text{Plan: } [\text{surface_desc_obj}(*O, *C, *R, *T)]). \end{aligned} \quad (4.16)$$

$$\begin{aligned} \text{Act}(\text{describe_event_type}(*E, *C, *At), \\ \text{Appcon: } *C = \{ \text{type}(*E, *T) \}, \\ \text{Effect: } \text{described_event}(*E, *C, *At)). \end{aligned} \quad (4.17)$$

$$\begin{aligned} \text{Decomp}(\text{describe_event_type}(*E, *C, *At), \\ \text{Plan: } [\text{surface_desc_event}(*E, *C, *At)]). \end{aligned} \quad (4.18)$$

図 4.6: Elaboration 関係に基づく発話プランニングのための諸定義

Elaboration 関係

図 4.6 に Elaboration 関係に基づいてドメイン行為の内容を段階的に記述するための行為スキーマと分解メソッドを示す．分解メソッド (4.14) において，記法 $F(X, Y, \dots) =.. [F, X, Y, \dots]$ は項 $F(X, Y, \dots)$ をリスト $[F, X, Y, \dots]$ に分解するために用いられる．これらの行為スキーマと分解メソッドを適用することによって，ドメイン行為を採用させるという発話プランは，ドメイン行為の各構成要素を別々に記述するという発話プランへ詳細化される．

表層的な対話行為 *surface_desc_obj* はドメイン対象を記述する言語表現を発話するために用いられる．このとき，協調的対話原則 8 にしたがって，注視されているドメイン対象は代名詞化する．また，協調的対話原則 9 にしたがって，注視状態に対して適切な発話を行うために提題化する必要がある対象は提題化する．

例にもどって，発話プラン (4.10) の最初の対話行為に図 4.6 の行為スキーマと分解メソッドを適用していくと次の発話プランが得られる．

$$\begin{aligned} & \text{surface_desc_obj}(x1, \{\text{type}(x1, \text{building}), \\ & \quad \text{named}(x1, \text{“武蔵野センタ”})\}, \\ & \quad \text{source, move}), \\ & \text{surface_desc_obj}(x2, \{\text{type}(x2, \text{bus})\}, \text{manner, move}), \\ & \text{surface_desc_obj}(x3, \{\text{type}(x3, \text{station}), \text{nearest}(x3, x1)\}, \\ & \quad \text{dest, move}), \\ & \text{surface_desc_event}(a1, \{\text{type}(a1, \text{move})\}, \text{proposal}). \end{aligned} \tag{4.19}$$

これは表層的な対話行為の列であり，発話実行部で実行可能である．詳細なドメインプラン (4.3) が立案される以前に，発話を開始することが可能であることが分かる．この表層的な対話行為の列によって生成される発話は 4.6.4 節で示す．

Circumstance 関係

図 4.7 に Circumstance 関係に基づく発話プランニングを実現するための行為スキーマと分解メソッドを示す．これらは，ドメイン事象を記述するときに，事象を構成する対象が存在するという前提条件を提示するために用いられる．述語

$$\text{Act}(\text{describe_obj_with_existential_precond}(*O, *C, *R, *E), \quad (4.20)$$

$$\text{Effect: described_obj}(*O, *C) \wedge \\ \text{described_thematic_rel}(*R, *O, *E)).$$

$$\text{Decomp}(\text{describe_obj_with_existential_precond}(*O, *C, *R, *E), \quad (4.21)$$

$$\text{Appcon: get_existential_precond}(*Precond, *O, *E), \\ \text{Plan: } [\text{describe_existential_precond}(*Precond, *O, *C), \\ \text{describe_obj_with_thematic_rel}(*O, *C, *R, *E)].$$

$$\text{Act}(\text{describe_existential_precond}(*Precond, *O, *C), \quad (4.22)$$

$$\text{Effect: described_obj}(*O, *C) \wedge \\ \text{described_thematic_rel}(\text{subj}, *O, *Precond) \wedge \\ \text{described_event}(*Precond, \\ \{type(*Precond, exist)\}, \\ \text{present_reason})).$$

$$\text{Decomp}(\text{describe_existential_precond}(*Precond, *O, *C), \quad (4.23)$$

$$\text{Plan: } [\text{surface_desc_obj}(*O, *C, \text{subj}, \text{exist}), \\ \text{surface_desc_event}(*Precond, \{type(*Precond, exist)\}, \\ \text{present_reason})].$$

図 4.7: Circumstance 関係に基づく発話プランニングのための諸定義

get_existential_precond はドメイン事象を構成する対象が存在するという前提条件を取り出すために使われる．話し手の態度 *present_reason* は，ある事象を他の事象の成立の理由として提示するという話し手の態度を表す．

ここで森の里青山行きというバスに乗るというドメイン行為 *a8* を提案することを考えよう．この行為の情報内容は次のように表される．

$$\begin{aligned} & \text{cont}(\text{a8}, \{\text{type}(\text{a8}, \text{geton}), \text{dest}(\text{a8}, \text{x8})\}). \\ & \text{cont}(\text{x8}, \{\text{type}(\text{x8}, \text{bus}), \text{named}(\text{x8}, \text{“森の里青山行き”})\}). \end{aligned} \quad (4.24)$$

このドメイン行為 *a8* を対話相手に採用させるという対話目標を達成するために図 4.6 で示した (4.11) ~ (4.14) を適用すると，次の発話プランが立案される．

$$\begin{aligned} & \text{achieve}(\text{described_obj}(\text{x8}, \{\text{type}(\text{x8}, \text{bus}), \\ & \quad \text{named}(\text{x8}, \text{“森の里青山行き”})\})), \\ & \text{achieve}(\text{described_thematic_rel}(\text{dest}, \text{x8}, \text{a8})), \\ & \text{achieve}(\text{described_event}(\text{a8}, \{\text{type}(\text{a8}, \text{geton})\}, \text{proposal})). \end{aligned} \quad (4.25)$$

ここで (4.25) の最初の 2 つの対話行為に図 4.7 の行為スキーマ (4.20) と分解メソッド (4.21) を適用し，その結果得られる対話行為の列に図 4.6 と図 4.7 の行為スキーマと分解メソッドを適用していくと，次の発話プランが得られる．

$$\begin{aligned} & \text{surface_desc_obj}(\text{x8}, \{\text{type}(\text{x8}, \text{bus}), \\ & \quad \text{named}(\text{x8}, \text{“森の里青山行き”})\}, \\ & \quad \text{subj}, \text{exist}), \\ & \text{surface_desc_event}(\text{e1}, \{\text{type}(\text{e1}, \text{exist})\}, \text{present_reason}), \\ & \text{surface_desc_obj}(\text{x8}, \{\text{type}(\text{x8}, \text{bus}), \\ & \quad \text{named}(\text{x8}, \text{“森の里青山行き”})\}, \\ & \quad \text{dest}, \text{geton}), \\ & \text{surface_desc_event}(\text{a8}, \{\text{type}(\text{a8}, \text{geton})\}, \text{proposal}). \end{aligned} \quad (4.26)$$

この表層的な対話行為の列によって生成される発話は 4.6.4 節で示す．

$$\text{Act}(\text{propose_act_with_motivation}(*A), \text{Effect: persuaded_act}(*A)). \quad (4.27)$$

$$\begin{aligned} \text{Decomp}(\text{propose_act_with_motivation}(*A), & \quad (4.28) \\ \text{Appcon: } \text{get_motivation}(*P, *A), & \\ \text{Plan: } [\text{propose_act}(*A), \text{assert_motivation}(*P, *A)]) & \end{aligned}$$

$$\text{Act}(\text{assert_motivation}(*P), \text{Effect: persuaded_prop}(*P)). \quad (4.29)$$

$$\begin{aligned} \text{Decomp}(\text{assert_motivation}(*P), & \quad (4.30) \\ \text{Appcon: } \text{cont}(*P, *C), & \\ \text{Plan: } [\text{described_event}(*P, *C, \text{present_reason})]) & \end{aligned}$$

図 4.8: Motivate 関係に基づく発話プランニングのための諸定義

Motivate 関係

図 4.8 に Motivate 関係に基づく発話プランニングを実現するための行為スキーマと分解メソッドを示す。これらは、ドメイン行為を提案するときに、対話相手がそのドメイン行為を採用することを動機付ける命題を提示するときに用いられる。述語 *get_motivation* はドメイン行為の採用を動機付ける命題を取り出すために使われる。たとえば、「愛甲石田駅で降りる」というドメイン行為を提案するときに、「愛甲石田駅が厚木研究センタの最寄駅である」という命題を動機付けのために提示するといった談話を生成することが考えられる。

4.6.4 発話実行

発話実行部は、発話プランニング部から送られる対話行為を言語表現として実現し、音声合成部に送る。このとき、各発話単位の間決められた一定のポーズを置くように、言語表現を音声合成部に送る。

発話単位としては名詞句を一つ含むような単位を使う。発話プランニング部から送られる表層的な対話行為が発話単位として小さすぎる場合があるので、次の発話単位統合ルールを用いる。

発話単位統合ルール 発話プランニング部から次の形をもつ 2 つの表層的な対話行為の列が送られてきた場合には，その 2 つの対話行為を一つの発話単位として生成せよ．

$$[surface_desc_obj(*O, *C, *R, *E), \\ surface_desc_event(*E, \{type(*E, *T)\}, *At)]$$

たとえば，発話プラン (4.19) に基づいて次の発話が実行される．’/’ は発話単位の区切りを表す．発話プラン (4.19) の 3 番目と 4 番目の対話行為は発話単位統合ルールによって一つの発話単位として実行されている．

談話 (d4.2)

(u4.2.1) 武蔵野センタからはですね / バスで / 最寄りの駅まで
行ってですね

発話プラン (4.26) に基づいて実行される発話は次のようになる．発話においてドメイン対象 x_8 が二度記述されるが，二度目に言語実現されるときには，協調的対話原則 8 にしたがって代名詞化されている．

談話 (d4.3)

(u4.3.1) 森の里青山行きのバスが ありますので / それに 乗って下さい

4.6.5 発話プランニングと発話実行の並行進行による漸次的発話生成

漸次的な発話生成は，問題解決，発話プランニング，発話実行を並行して進めることによって実現される．抽象的なドメインプランに基づいて発話プランが立案され，その発話プランに基づいて発話を実行している間に，より詳細なドメインプランが得られると，その新しいドメインプランに基づいて新しい発話プランの立案が始まる．新しい発話プランの立案と古い発話プランに基づく発話実行は，並行して進められる．新しい発話プランに基づく発話の実行が可能になった時点

で，古い発話プランに基づく発話実行は中断され，新しい発話プランに基づく発話を開始される．

例えば，詳細なドメインプラン (4.3) が立案される以前に，抽象的なドメインプラン (4.1) に基づいて立案された発話プラン (4.19) にしたがって，発話 (u4.2.1) が実行されている状況を想定する．

この発話が行われている間にドメインプランは (4.1) から (4.3) に詳細化されていく．詳細化されたドメインプラン (4.3) が完成すると，次の新しい発話プランが立案される．

```
surface_desc_obj(x1, {type(x1, building),  
                      named(x1, “武蔵野センタ”)}},  
                 source, move),  
surface_desc_obj(x2, {type(x2, bus)}},  
                 manner, move),  
surface_desc_obj(x7, {type(x7, station),  
                      named(x7, “吉祥寺”)}},  
                 dest, move),  
surface_desc_event(a4, {type(a4, move)}, proposal).
```

(4.31)

古い発話プランに基づく発話 (u4.2.1) の「最寄りの駅まで」が発話された時点で，新たな発話プラン (4.31) が得られたとする．この時点で発話 (u4.2.1) は中断され，発話プラン (4.31) に基づく発話が再開される．このとき，協調的対話原則 7 によって既に伝達済みの情報は繰り返されない．結局，次のような発話が生成されることになる．

談話 (d4.4)

(u4.4.1) 武蔵野センタからはですね / バスで / 最寄りの駅まで /
吉祥寺まで 行ってですね /

発話 (u4.4.1) においては「最寄りの駅まで」という発話が後続の「吉祥寺まで」という発話によって言い直されたとみなすことができる．

このようにして，発話内容を組み立てながら発話を実行し，発話を実行しながら発話内容を詳細化することが可能となり，システムが発話すべきときに即座に発話を開始することが容易となる．生成される談話の適切性は協調的対話原則によって保証される．

4.6.6 対話相手からの応答に対処するための対話制御

発話生成部は，話し言葉に特有の小さな発話単位を使って段階的に情報をユーザに伝達しながら，協調的対話原則にしたがって発話理解部，対話制御部と連動することにより，ユーザからのアイツチや割り込み発話に臨機応変に対処する．

発話理解部は，常時ユーザ発話を受け付け，対話状態にしたがってユーザ発話の応答タイプを分類する．ここでは，システム発話中に起きるユーザ発話の応答タイプの代表例として，次の応答タイプを取り上げる．

Accept: ユーザが現在のシステム発話を承認した．

Reject: ユーザが現在のシステム発話の承認を拒否した．

Confirm(NP): ユーザが現在のシステム発話に含まれる名詞句 NP について確認した．

Request(NP): ユーザが現在のシステム発話の伝達内容以外の情報として，名詞句 NP に関連する情報を要求した．

Accept はアイツチ，Confirm(NP) は名詞の復唱による確認発話，Reject と Request(NP) は割り込み発話に対応する．一般的な応答解析の実現は本章の目的ではないので，ここでは簡易な応答解析を想定する．ユーザが発話する言語表現としては，肯定的な間投的表現 (例: 「はい」)，承認を意味する表現 (例: 「分かりました」)，拒否を意味する表現 (例: 「分かりません」)，名詞句のみを考える．ここで強調しておきたいことは，応答タイプは言語表現だけでは決まらず，対話状態という文脈に依存して決定されるということである．次に応答タイプ決定ルールを示す．

応答タイプ決定ルール

If システム発話に対して，ユーザが肯定的な間投的表現あるいは，承認を意味する表現を発話した

Then ユーザ発話の応答タイプは Accept である

ElseIf システム発話に対して，ユーザが拒否を意味する表現を発話した

Then ユーザ発話の応答タイプは Reject である

ElseIf システム発話に対して，ユーザがシステム発話に含まれる名詞句，あるいは，名詞句の一部を発話した

Then ユーザ発話の応答タイプは Confirm(NP) である

ElseIf システム発話に対して，ユーザがシステムに発話に含まれない名詞句を発話した

Then ユーザ発話の応答タイプは Request(NP) である

対話制御部は，ユーザ応答が入力された場合だけでなく，協調的対話原則 4 にしたがって，新情報を担う節単位に対してユーザが応答しない場合も考慮する必要がある．対話制御部は次のように動作する．

対話制御ルール

- ユーザ発話の応答タイプが Accept である場合:
 1. 発話理解部は，現在のシステム発話に対応する対話行為がユーザに承認されたことを対話状態に書き込む．
 2. 協調的対話原則 5 にしたがって，発話は現在の発話プラン通りに進める．
- ユーザ発話の応答タイプが Reject であるか，あるいは，システムが新情報を担う節単位の発話を実行した後にユーザからの応答がない場合:
 1. 対話状態を参照して，ユーザに承認することを拒否された対話行為を同定する．
 2. 発話プランニング部に対して，協調的対話原則 4 にしたがって，拒否された対話行為を詳細化，あるいは，その対話行為の核または衛生を記述するような発話プランを立案するように命じる．

3. 発話プランニング部は、命じられた通りの発話プランが立案可能であるなら、現在の発話実行を停止するように発話実行部に命じ、発話プランの立案を開始する。命じられた通りの発話プランの立案ができないなら、現在の発話を続行する。

- ユーザ発話の応答タイプが Cofirm(NP) である場合:

1. 現在の発話実行を中断するように発話プランニングと発話制御部に命じる。
2. 協調的対話原則 6 にしたがって、発話プランニング部に対してユーザの復唱応答に対処するための発話を生成するように命じる。
3. 中断していた発話を再開するように、発話プランニング部と発話制御部に命じる。

- ユーザ発話の応答タイプが Request(NP) である場合:

1. 現在の発話実行を中断するように発話プランニングと発話制御部に命じる。
2. 発話プランニング部に対して、名詞句 NP で示される情報を先に伝達するような発話プランを最立案するように命じる。

ここでは、便宜上、ユーザ応答のタイプが Reject の場合は、新情報を担う節単位に対してユーザの応答がない場合と同等に扱っている。いずれの場合も、システムが実行した対話行為が所定の対話目標を達成できなかった場合であり、協調的対話原則 4 にしたがって、失敗した対話行為を詳細化するか、その対話行為の核または衛生を記述するような発話を生成し、ユーザからの承認を得る機会を待つ。また、ユーザ応答タイプが Request(NP) であるときは、協調的対話原則とは別に、ユーザが知りたがっている情報を先に伝達するように話の流れを変更する。

4.7 考察

提案法を実装した実働システムが生成した対話を例にとりながら、提案法と協調的対話原則の有効性について考察する。交通経路案内タスクにおいて、提案法

談話 (d4.5)

- (u4.5.1) 武蔵野センタからは / バスで / 最寄りの駅まで /
- (u4.5.2) えーと 吉祥寺までですね / 出て下さい /
- (u4.5.3) で / 井の頭線で / 下北沢まで 行って /
- (u4.5.4) 小田急線で / 愛甲石田に 行って下さい /
- (u4.5.5) それから / バスなのですが /
- (u4.5.6) 北口バス乗り場というバス乗り場が ありますので /
- (u4.5.7) そこで / 森の里青山行きというバスに 乗って /
- (u4.5.8) 森の里青山行きのバスが ありますので /
- (u4.5.9) 通信研究所前で 降ります /

図 4.9: システムが生成した談話 (d4.5)

談話 (d4.6)

- (u4.6.1) それで / 小田急線の愛甲石田まで 行けばよいのですが
- (u4.6.2) えーと 吉祥寺からはですね / 井の頭線で /
下北沢まで 行って下さい /

図 4.10: システムが生成した談話 (d4.6)

を Common Lisp を使って実装した．問題解決部と発話プランニング部の実装には論理制約単一化システム [63] を利用した．問題解決部を実装するために，18 個の行為スキーマと 16 個の分解メソッドを記述した．発話プランニング部については，4.6 節で示したものを含めて 16 個の行為スキーマと 16 個の分解メソッドを記述した．システムには，120 の駅と建物，25 の路線を含む地図の上での交通経路探索問題を与えた．図 4.4 で示した地図は実験で使った地図の一部である．システム発話途中のユーザ応答の認識は単語認識によって行った．認識語彙は約 50 単語であり，問題ごとに認識語彙は変更した．

まず，4.6.5 節で説明した発話プランニングと発話実行の並行進行による漸次的発話生成という観点から，提案法と協調的対話原則の有効性について考察する．ここでは，生成される談話の適切性を保証するために用いられる協調的対話原則 7，協調的対話原則 8，協調的対話原則 9 に着目する．提案方法と協調的対話原則の有効性について考察するために，協調的対話原則を利用した場合と利用しなかった場合とで生成された談話を比較した．

協調的対話原則を利用した場合にシステムは適切な談話を生成した．図 4.9 に武蔵野センタから厚木センタまで行くという問題を与えたときにシステムが生成した談話を示す．4.6.5 節で説明した通り，発話 (u4.5.1) の発話と並行して，問題解決と発話プランニングは進められている．発話 (u4.5.2) の冒頭の言い淀み「えーと」は時間制限を守るために発話された．発話 (u4.5.2) では，4.6.5 節で述べたように，協調的対話原則 7 が適用され，冗長な発話が避けられている．協調的対話原則 7 を使わないと，発話 (u4.5.1) で発話された「武蔵野センタからは」，「バスで」といった表現が発話 (u4.5.2) でも繰り返され，冗長で不自然な談話が生成された．発話 (u4.5.2) の終了時点で既に吉祥寺から愛甲石田までの経路が決定されており，発話 (u4.5.3)，(u4.5.4) が発話された．また，協調的対話原則 8 は主として移動行為の源泉格のゼロ代名詞化のために用いられている．

発話 (u4.5.2) の終了時点で吉祥寺から目的地の最寄駅である愛甲石田の間の経路が詳しく決定されておらず，吉祥寺から愛甲石田まで行くという内容のドメインプランしか得られていない場合には，発話 (u4.5.2) に続いて図 4.10 に示す談話が生成された．発話 (u4.6.1) の終了時には愛甲石田が注視されている．発話 (u4.6.1) が終了したとき，吉祥寺から愛甲石田まで経路が詳細化されており，発話 (u4.6.2)

対話 (d4.7)

- (u4.7.1) システム: 小田急線に乗っ< はい> て /
- (u4.7.2) システム: 愛甲石田で降りてください /
- (u4.7.3) システム: 愛甲石田が / NTT 厚木センタの最寄りの駅になりますので /
- (u4.7.4) ユーザ: はい
- (u4.7.5) システム: そこで降りて /
- (u4.7.6) システム: 北口バス乗り場というバス乗り場が
- (u4.7.7) ユーザ: 北口バス乗り場
- (u4.7.8) システム: はい ありますので /
- (u4.7.9) システム: そこで / 森の里青山行きというバスに
- (u4.7.10) ユーザ: 森の里
- (u4.7.11) システム: 青山行き / < はい> それに乗って /

図 4.11: ユーザのアイツチや復唱へ対処しながらの対話 (d4.7)

からその経路を説明し始めるが、協調的対話原則 9 により、吉祥寺からの経路を説明するためには注視状態を愛甲石田から吉祥寺に戻す必要がある。したがって、吉祥寺を提題化した「吉祥寺からはですね」という表現がまず発話された。協調的対話原則を適用しない場合には、冗長で、注視状態に対して不適切な発話が生成された。

次に、4.6.5 節で説明したユーザのアイツチや割り込み発話に対処しながら発話を生成するという観点から、提案法と協調的対話原則の有効性について考察する。ここでは、対話相手からの応答の有無に対処するために用いられる協調的対話原則 2、協調的対話原則 3、協調的対話原則 4、協調的対話原則 5、協調的対話原則 6

に着目する．図 4.11 にシステムがユーザからのアイツチや名詞の復唱に対処しながら行った対話を示す．

システムはユーザからのアイツチ，復唱に対して臨機応変に対処できた．これは，(i) 協調的対話原則 1 にしたがって，システムは話し言葉特有の小さな発話単位を使って段階的に情報をユーザに伝達するため，システム発話途中のユーザのアイツチや割り込みに即座に対処できる，(ii) 発話内容の組み立てと発話の実行を並行して進めることにより，システムが発話すべきときに即座に発話を開始することが容易である，(iii) 協調的対話原則 5，協調的対話原則 6 にしたがって，発話生成部が発話理解部，対話制御部と連動することにより，システムはユーザからのアイツチや復唱に適切な戦略に対処しながら，発話を生成できるという理由による．

さらに重要なことには，協調的対話原則 2，協調的対話原則 3，協調的対話原則 4 にしたがうことにより，ユーザから応答がなくてもシステムが談話戦略を変更する場合があるということである．図 4.11 の対話において，発話 (u4.7.2) は新情報を担う節単位であり，この発話に対してユーザが応答しなかった．協調的対話原則 3，協調的対話原則 4 にしたがうことにより，システムは，その時点で計画済みの発話 (u4.7.6) の実行を中断して，発話 (u4.7.2) の衛星となる発話 (u4.7.3) を立案し実行した．発話 (u4.7.3) は愛甲石田で降りるという行為を採用するように対話相手を動機付けるための発話である．この発話は，図 4.8 で示した Motivate 関係に基づく発話プランニングのための行為スキーマ，分解メソッドによって立案された．これらの協調的対話原則がないと，システムはユーザの不理解または未承認を含意する沈黙に適切に対処できない．

発話 (u4.7.4) では，ユーザが発話 (u4.7.3) に対して肯定的な間投的表現「はい」を発話したので，システムは協調的対話原則 4 にしたがうことにより，中断していた発話発話 (u4.7.6) を再開しようとした．協調的対話原則 9 により，中断していた発話を新たな文脈で再開するためには，発話を中断したときの注視状態に戻す必要がある．そこで，発話 (u4.7.6) を生成する前に，発話 (u4.7.2) と同じ内容をもつ発話 (u4.7.5) が生成された．ただし，発話 (u4.7.5) では「愛甲石田」は代名詞化されている．

発話 (u4.7.5) に対してはユーザからの応答はないが，発話 (u4.7.5) は旧情報を

対話 (d4.8)

- (u4.8.1) システム: 武蔵野センタからは / バスで /
- (u4.8.2) ユーザ: 愛甲石田からは
- (u4.8.3) システム: はい / 小田急線で / 愛甲石田に 行って下さい / < はい >
- (u4.8.4) システム: それから / バスなんですが /

図 4.12: ユーザの割り込み発話に対処しながらの対話 (d4.8)

担う節単位であるので、システムは、協調的対話原則 2 にしたがうことにより発話を継続し、発話 (u4.7.6) を生成した。この協調的対話原則がないと、システムはユーザの沈黙に過剰に反応して不必要に冗長な発話を生成することになる。

続いて、図 4.12 にシステムがシステムからの割り込み発話に応じて話の流れを変更した対話を示す。システム発話 (u4.8.1) の途中で、ユーザはシステム発話に含まれない名詞「愛甲石田」を発話することによって、割り込み発話を行っている。4.6.6 節で説明した対話制御ルールにより、システムは、愛甲石田を含む経路を先に伝達するように、発話プランを再立案し、発話 (u4.8.3) で割り込み発話に対処する発話を始めている。提案法は、話し言葉特有の小さな発話単位を使っているので、ユーザの割り込み発話に即座に対処できている。

以上の考察で分かるように、提案法は、協調的対話原則にしたがうことによって、システムがユーザからのアイツチや割り込み発話に即座に対処し、臨機応変に話の進め方を変更しながら、自然な発話を生成することを可能とする。

4.8 おわりに

本章では、音声対話システムが、ユーザからのアイツチや割り込み発話に即座に対処し、臨機応変に話の進め方を変更しながら、自然な発話を生成することを可能とする漸次的発話生成法を示した。提案法は一連の発話を一つの談話として生成する。漸次的に生成される談話の適切性と、システム発話途中に起きるユーザ発話に適切に対処できることは、人同士の対話を書き起こした対話データの分

析と言語学的な知見から導いた協調的対話原則によって保証される．提案法は以下の利点をもつ．

- システムは話し言葉特有の小さな発話単位を使って段階的に情報を伝達するので，システム発話途中のユーザからのアイツチ，復唱発話，割り込み発話に即座に対処することができる．提案法は，主導権混在型音声対話システムのシステム対応フェーズにおいて，従来法では実現できない自由な主導権交代を可能とする．
- 提案法は，発話内容の組み立てと並行して発話を実行し，発話内容が部分的に決定した時点で，話し言葉特有の小さな発話単位を使って発話を開始することができるので，システムが発話すべきときに即座に発話を開始することが容易となる．
- 提案法は，協調的対話原則にしたがって，発話理解部，発話生成部，対話制御部が適切に連動することにより，ユーザからのアイツチ，復唱発話，割り込み発話，沈黙に即座に対処し，ユーザの意図に応じて臨機応変に話の進め方を変更することができる．このことにより，システムは，対話の中でやり取りされる情報についてユーザとの間で共通理解を維持しながら対話を進めることが可能となる．

なお，提案法は，本章で例にとった交通経路案内タスクだけでなく，テレビ番組録画予約タスク，天気情報案内タスクにおいても実装され，実働する音声対話システムに組み込まれている [28, 64] ．

今後の課題としては以下のことがある．

- 本研究では，交通経路案内タスクの対話データを分析した結果から協調的対話原則を導いたが，いかなるタスクでも成立するような一般的な原則を導くためには，複数タスクの対話データの分析に基づいて協調的対話原則を導く必要がある．
- 本研究では，音声対話システムの行動指針を与える協調的対話原則を人同士の対話を書き起こした対話データから導いた．このことにより，音声対話に

おける発話生成の工学的モデルを構築するために役立つ知見が得られたと考える。しかし、人とシステムの対話は、人と人の対話と同等ではなく、独自の特徴のもつはずである。今後は、システムとユーザの対話データから協調的対話原則を獲得し、人とシステムの対話に即した行動指針を導くことが考えられる。

- 本研究で導いた協調的対話原則は、対話の状態におけるシステムの振る舞いを決定論的に記述している。しかし、ユーザの振る舞いは確率論的であって、それに対するシステムの振る舞いを決定的に記述することは、システムの振る舞いの柔軟さを損ねる可能性があり、必ずしもシステムの最適な振る舞いを実現することにつながるとは限らない。実際、マルコフ決定過程上でのプランニングにおいては、確率的な行動指針が最適となる場合があることが知られている。今後は、協調的対話原則を各対話状態における確率的な行動指針として与えることによって、より柔軟で最適なシステムの振る舞いを実現することが考えられる。
- 本研究では、提案法によってユーザのアイツチや割り込み発話に即座に対処しながら自然な発話を生成することが可能となり、このことは書き言葉の文を単位として発話を生成する従来法では実現が難しいという定性的な議論を行った。今後は、タスク達成率、対話の効率、ユーザ満足度といった評価尺度を使って、本研究で提案した漸次的発話生成法と協調的対話原則の効果を定量的に評価していくことが必要である。

第 5 章

効率的対話制御

5.1 はじめに

音声対話システムがユーザと円滑に意思疎通を行うためには、システムはユーザ音声の認識誤りに対処する必要がある。本章では、ユーザがシステムのデータベース内容についての問い合わせを行い、システムが問い合わせ内容に応じてユーザが必要とする情報を伝達するという情報案内タスクに焦点をあてる。そのようなタスクを遂行するシステムは、ユーザ音声の認識結果からユーザ問い合わせ内容を理解し、問い合わせ内容に応じて適切な情報をユーザに伝達しなければならないが、ユーザ音声の認識誤りの可能性があるため、認識結果のみに頼ってユーザ問い合わせ内容を決定してしまうと、ユーザが必要とする情報を伝達できない場合が生じる。

主導権混在型の音声対話システムでは、この問題に対処するために、ユーザとの間で確認対話と呼ぶ一連のやり取りを行うことが多い。確認対話においては、システムは認識した結果をユーザに問い返すことによって確認し、ユーザは確認内容が正しければ、承認発話（例：「はい」）により確認内容を承認し、確認内容が正しくないなら訂正発話によって確認内容を訂正する。ユーザが承認した問い合わせ内容は承認済みとみなされる。確認対話が終了した後、システムはその時点で理解している問い合わせ内容にしたがって、適切な情報を伝達するためにシステム応答を生成する。

確認対話は音声認識誤りに対処するための有効な方略ではある。しかし、確認

対話 (d5.1)

- (u5.1.1) ユーザ: 神奈川県に大雨警報が発表されていますか？
- (u5.1.2) システム: 香川県の大雨警報についてですか？
- (u5.1.3) ユーザ: 神奈川県です．
- (u5.1.4) システム: 香川県の大雨警報についてですか？
- (u5.1.5) ユーザ: 神奈川県です．
- (u5.1.6) システム: 神奈川県の大雨警報についてですか？
- (u5.1.7) ユーザ: はい
- (u5.1.8) システム: 神奈川県に大雨警報は発表されていません．

図 5.1: 不必要な確認を伴う非効率な対話の例

対話の最中にも音声認識誤りが起きる可能性があり，確認対話が長ければ長いほど，対話の円滑な流れが阻害される危険性が高まる．したがって，不必要な確認はできる限り避けることが望ましい．本研究では，不必要な確認を省き，できるだけ短い対話でユーザが必要とする情報を伝達することを可能とする対話制御法に焦点をあてる．対話制御とは，対話の各時点においてシステムの行動を決定する機能であり，本研究は対話の長さを最小にするという観点から対話の各時点における最適なシステム行動を決定する対話制御法に着目する．

不必要な確認が生じる典型的な状況の一つは，ユーザがシステムのデータベース内容の範囲を超えた問い合わせを行うときや，データベースの内容では答えることができないほど詳細な情報を求めるという状況である．このとき，システムが問い合わせ内容の一部を把握しさえすれば，問い合わせには応えることができないと伝えたり，システムにとって提供できるだけの情報を提供することができるにもかかわらず，問い合わせ内容をすべて逐一確認しようとするすると，不必要な確認が生じることとなる．

天気情報案内システムを例にとって説明する．システムのデータベースには，各

対話 (d5.2)

- (u5.2.1) ユーザ: 神奈川県に大雨警報が発表されていますか？
- (u5.2.2) システム: 警報についてですか？
- (u5.2.3) ユーザ: はい
- (u5.2.4) システム: 警報はどこにも発表されていません．

図 5.2: 不必要な確認を避けた効率的な対話の例

地の今日、明日の天気、気温、降水確率、現在発表されている警報についてのデータが記録されている．ここで、警報はどこにも発表されていないことがデータベースに記録されているときに、ユーザが神奈川県に大雨警報が発表されているかどうか問い合わせたという状況を想定する．この状況の下で、図 5.1 に示す対話では、システムは場所、警報の種類、情報の種別について確認している．システムは神奈川県を香川県と誤認識しており、確認の最中にもその誤認識が繰り返し起きている．

しかし、警報がどこにも発表されていないという現在のデータベース内容から判断すれば、システムは場所や警報の種類について確認する必要はないことが分かる．なぜなら、図 5.2 に示すように、それらの確認がなくても、警報という情報種別だけを確認すれば、システムはどこにも警報は発表されていないという情報をユーザに伝達することが可能であり、場所や警報種別についての確認をした場合と比較して、システムの応答が長くなるわけではないからである．ただし、多くの場所で警報が発表されているといった別の状況では、場所について確認する必要がある．なぜなら、場所についての確認を省略すると、システムは警報が発表されている場所をすべて列挙することになり、システム応答が極端に長くなってしまうからである．

このように、不必要な確認を避けるためには、確認のためのやり取りだけでなく、確認後のシステム応答の長さも含めて、ユーザに必要な情報を伝達するまでの対話全体の長さを考慮する必要がある．すなわち、確認が必要かどうかを対話制御部が決定するためには、その時点のシステム理解状態だけでなく、確認後の

システム応答の長さを発話生成部に問い合わせることが必要であり，対話制御部は，発話理解部，発話生成部と適切に連動しなければならない．

Walker 等は音声対話システムの有用性 (ユーザビリティ) を計る尺度として，ユーザ満足度を使うべきであり，さらに，ユーザ満足度はタスク達成率や対話の効率など他の尺度から予測することが可能であることを論じている [84]．本章は，できるだけ短い対話でユーザが必要とする情報を伝達するという意味で，効率の良い対話でタスクを達成するための対話制御に関心があり，ユーザ満足度については考察しない．ユーザ満足度は重要な尺度ではあるが，本研究は，対話の効率，タスク達成率という客観的な尺度に着目して，数理的に定式化可能なモデルを構築することにより，音声対話システムの有用性を計るという目的に対しても貢献するものと考えている．

音声対話システムとユーザの間で効率的な対話を実現するための対話制御法が従来より提案されてきている [11, 50, 53, 66, 75]．これらの従来法は，音声認識結果の信頼度，音声認識精度といった情報を利用して，確認対話の長さを削減することに注目し，確認対話に続くシステム応答の長さを含めて対話全体の長さを削減することを十分に考慮していないため，これらの従来法を直接適用しても，データベース内容に依存する不必要な確認を避けることは難しい．

音声認識結果の信頼度が閾値を超えているならば，確認を行わない対話制御法が提案されている [47]．この方法は理解誤りを最小にするという目的の下で確認数を削減するための対話制御法とみなすことができる．一方，本研究では，対話全体の長さを最短にするという目的の下での対話制御に着目し，対話全体の長さを最短にすることによって，結果的に不必要な確認を回避できることを示す．

ユーザ発話内容が曖昧なときに，ユーザ発話内容の曖昧さを解消してもシステム応答が同一で変化しないなら，ユーザ発話内容の曖昧さを解消せずに応答を生成するという方法が提案されている [2, 71, 83]．しかし，これらの従来法は，システム応答の同一性が保証されていない場合には適用できないという問題がある．また，音声認識誤りにより発生する対話については考慮されていない．

本研究では，従来法の問題点を解決するために，システム応答を含めて対話全体の長さを最小にするように，対話各時点で最適なシステムの行動を選択する対話制御法として，デュアルコスト法を提案する．デュアルコスト法は，確認対話の

長さ (確認コスト) と確認対話後にユーザに情報を伝達するために生成されるシステム応答の長さ (情報伝達コスト) の和を対話コストと捉え、対話コストを最小化するように対話を制御する。確認コストは音声認識精度に依存し、情報伝達コストはシステムのデータベースの内容に依存する。確認コストと情報伝達コストは相反する関係にある。デュアルコスト法は、確認対話だけでなくシステム応答を含めて対話全体の長さを最小化することにより、従来法では回避できない不必要な確認を避け、効率的な対話を実現できる。

確認対話が終了した後、ユーザに情報を伝達するために応答を生成する際の対話戦略として、確認対話で承認済みとなった情報のみを正しいと仮定して応答を生成する確定型の対話戦略と、未承認の情報も正しいと仮定して応答を生成する試行型の対話戦略を考えることができるが、本章では、確定型の対話戦略のみ考慮する対話制御法としてデュアルコスト法を提示し、その後、デュアルコスト法を発展させ、試行型の対話戦略も考慮する対話制御法として試行型デュアルコスト法を提案する。

確定型の対話戦略は確実であるが、確認対話が長くなる可能性がある。試行型の対話戦略は、確認対話を短くすることが可能だが、誤認識のため間違った情報伝達を行う危険がある。このときシステム応答は失敗する。しかし、ユーザはシステム応答が失敗したことを気づくことができ、必要な情報を手に入れるまで対話を継続すると仮定するならば、いったんは応答が失敗したとしても、何回か対話を繰り返しているうちに、ユーザが必要とする情報が伝達される確率は向上する。音声認識精度が比較的高い状況においては、試行型の対話戦略が確定型の対話戦略よりも平均的に短い対話でユーザに必要な情報を伝達できる場合があるはずである。

以下においては、デュアルコスト法、試行型デュアルコスト法について説明し、シミュレーション対話実験と被験者との対話実験の結果に基づいて、デュアルコスト法、試行型デュアルコスト法が従来法よりも短い対話で必要な情報を伝達できることを論じる。

5.2 音声対話システムにおける対話制御

ここでは、音声対話システムとユーザとの間の対話の流れについて述べ、音声対話システムにおける対話制御の役割について説明する。第2章で説明したように、情報案内タスクの音声対話システムとユーザの対話は、ユーザ問い合わせ把握フェーズとシステム情報伝達フェーズという2つのフェーズを移行しながら進行する。ユーザ問い合わせ把握フェーズでは、まず、ユーザがシステムに対する問い合わせを伝える。システムとユーザは確認対話を通してユーザ問い合わせ内容を把握する。システムが把握したユーザ問い合わせ内容はシステム理解状態として保持される。システム情報伝達フェーズでは、その時点でシステムが把握している問い合わせ内容に応じて、ユーザが必要とする情報を伝達するためにシステム応答が生成される。

ユーザ問い合わせ内容に対するシステム理解状態は、3つ組〈属性、値、承認フラグ〉の集合として保持される。タスクごとにユーザ問い合わせタイプの種類と属性の全体集合が与えられる。各ユーザ問い合わせタイプについて、ユーザ問い合わせ内容として含むことができる属性と、各属性がとりうる値の範囲が決まっている。各ユーザ問い合わせタイプについて、ユーザ問い合わせ内容として含むことができる属性を有効な属性、値としてとりうる属性値を有効な属性値とよぶ。各ユーザ問い合わせタイプについて、有効でない属性、属性値をそれぞれ無効な属性、無効な属性値とよぶ。承認フラグは、属性の値がユーザの承認発話によって承認されるまで「未」という値をとり、承認されると「済」という値をとる。

ユーザ問い合わせ把握フェーズでは、システムは確認行為、情報要求行為のいずれかの対話行為を遂行する。確認行為は、システム理解状態において値が与えられている属性について、値が正しいかどうかユーザに対して確認するための確認発話を行い、ユーザの承認発話によって値が承認されるまで、その属性の値の確認を繰り返すという行為である。ユーザは、システムの確認内容が誤っている場合には、システム理解状態を訂正するための訂正発話を行うことができる。情報要求行為とは、システム理解状態において値が与えられていない属性について、ユーザから属性の値を引き出すための情報要求発話を行い、その後、その属性についての確認行為を実施するという行為である。

システム情報伝達フェーズでは、システムはユーザが必要とする情報を伝達するためにシステム応答を生成するという情報伝達行為を行う。システムは、情報伝達行為を行なうとき、いくつかの属性値を正しいと仮定した上で応答文を生成する。それらの属性値を情報伝達行為や応答文の基礎となる属性値と呼ぶ。情報伝達行為やシステム応答が確定型であるとは、その基礎となる属性値が承認済みの属性値ばかりであることを言う。情報伝達行為やシステム応答が試行型であるとは、その基礎となる属性値の中に未承認の値が含まれることを言う。

対話制御とは、対話の各時点において、適切なシステムの対話行為を決定することである。本研究では、ユーザが必要とする情報をできるだけ短い対話で伝達するという効率性の基準に基づいて対話各時点のシステムの対話行為を決定する。システムの対話行為としては、確認行為、情報要求行為、情報伝達行為がある。情報伝達行為には、確定型の情報伝達行為と試行型の情報伝達行為がある。

本章では例として天気情報案内システムを取りあげる。システムのデータベースには、各場所について、今日、明日の天気、気温、降水確率と、現在発表されている警報のデータが記録されている。ユーザ問い合わせタイプとしては、警報問い合わせ、天気問い合わせ、気温問い合わせ、降水確率問い合わせの4種類の問い合わせタイプがある。

システム理解状態は、場所、日、警報種別、情報種別の4つの属性によって表現される。ユーザ問い合わせタイプと属性との間の関係を表5.1に示す。各ユーザ問い合わせタイプについて、有効な属性を○、無効な属性を×で示している。

場所属性は、いずれのユーザ問い合わせタイプでも有効な属性であり、特定の

表 5.1: 天気情報案内タスクにおけるユーザ問い合わせタイプと属性の関係

ユーザ問い合わせタイプ	属性			
	場所	日	警報種別	情報種別
天気問い合わせ	○	○	×	(値は天気のみ)
気温問い合わせ	○	○	×	(値は気温のみ)
降水確率問い合わせ	○	○	×	(値は降水確率のみ)
警報問い合わせ	○	×	○	(値は警報のみ)

場所の名前を値としてとる．日属性は，ユーザ問い合わせタイプが天気，気温，降水確率の問い合わせであるなら，今日か明日という値をとる．警報問い合わせでは，日属性は無効な属性である．警報種別属性は，大雨，洪水といった値をとり，天気問い合わせ，気温問い合わせ，降水確率問い合わせの各ユーザ問い合わせタイプにとっては無効な属性である．情報種別属性は，ユーザ問い合わせタイプに応じて，警報，天気，気温，降水確率という値をとる．警報という属性値は，ユーザ問い合わせタイプが警報問い合わせであるときにのみ有効であり，他のユーザ問い合わせタイプでは無効である．天気，気温，降水確率という属性値についても同様である．

図 5.2 で示した対話 (d5.2) を例にとる．この対話では，警報がどこにも発表されていないことがシステムのデータベースに記録されている．ユーザ問い合わせ把握フェーズにおいて，ユーザは発話 (u5.2.1) で問い合わせ内容をシステムに伝達している．システムが問い合わせ内容を正しく認識できたとすると，発話 (u5.2.1) を理解した後のシステム理解状態は次のようになる．

$$S_1 = \{ \langle \text{場所, 神奈川県, 未} \rangle, \langle \text{警報種別, 大雨, 未} \rangle, \quad (5.1) \\ \langle \text{情報種別, 警報, 未} \rangle \}$$

次に，システムは発話 (u5.2.2) で情報種別の値を承認済みとするための確認行為を行っている．ユーザは発話 (u5.2.3) でシステムの確認行為に対して承認発話を行っている．システムがこの承認発話を正しく認識できたとすると，システム理解状態は次のようになる．

$$S_2 = \{ \langle \text{場所, 神奈川県, 未} \rangle, \langle \text{警報種別, 大雨, 未} \rangle, \quad (5.2) \\ \langle \text{情報種別, 警報, 済} \rangle \}$$

続いて，システムは発話 (u5.2.4) で承認済みの情報種別の値のみを正しいとして，確定型の情報伝達行為を実行している．

図 5.3 は試行型の情報伝達行為の例を示している．ユーザが発話 (u5.3.1) で問い合わせ内容を伝達した後，システム理解状態は，対話 (d5.3) の場合と同じく，(5.1) のようになっているとする．このとき，システムは確認を一切行うことなしに，発話 (u5.3.2) で承認済みでない情報種別属性の値を正しいと仮定して試行型の情報伝達行為を実行している．

対話 (d5.3)

- (u5.3.1) ユーザ: 神奈川県に大雨警報が発表されていますか？
- (u5.3.2) システム: 警報はどこにも発表されていません。
{ 警報が正しいと仮定して，試行的にシステム応答を生成 }

図 5.3: 試行型の情報伝達行為の例

5.3 デュアルコスト法

5.3.1 概観

デュアルコスト法は，確認コストと情報伝達コストの和を最小化するように，対話の各時点におけるシステムの対話行為を選択する．確認コストと情報伝達コストの和を対話コストと呼ぶ．確認コスト，情報伝達コストという2種類のコストを導入するのは，5.1 節で述べたように，できるだけ短い対話でユーザが必要とする情報を伝達するためには，確認対話の長さだけでなく，それに続くシステム応答の長さも考慮する必要があるためである．確認コストと情報伝達コストは相反する関係にあり，2つのコストの和を最小にするような手順で対話を進めることにより，できるだけ短い対話でユーザが必要とする情報を伝達できる．

確認コスト，情報伝達コストは，対話の長さを計るために導入した概念である．対話の長さのどのように定義するかについては様々な方法を考えることができるが，ここでは，対話の長さをシステムとユーザがやり取りする自立語の数の期待値として定義する．すなわち，確認コストは確認対話においてシステムとユーザがやり取りする自立語の数の期待値であると定義し，情報伝達コストは，データベース内容をユーザ伝達するために生成されるシステム応答に含まれる自立語の数の期待値であると定義する．

デュアルコスト法におけるシステムの構成を図 5.4 に示す．システムは，発話理解部，対話制御部，発話生成部の3つの構成部に加えて，対話状態，データベースから構成される．対話状態にはシステム理解状態と，直前に実行したシステムの対話行為が記録されている．各構成部は対話状態を参照・更新する．

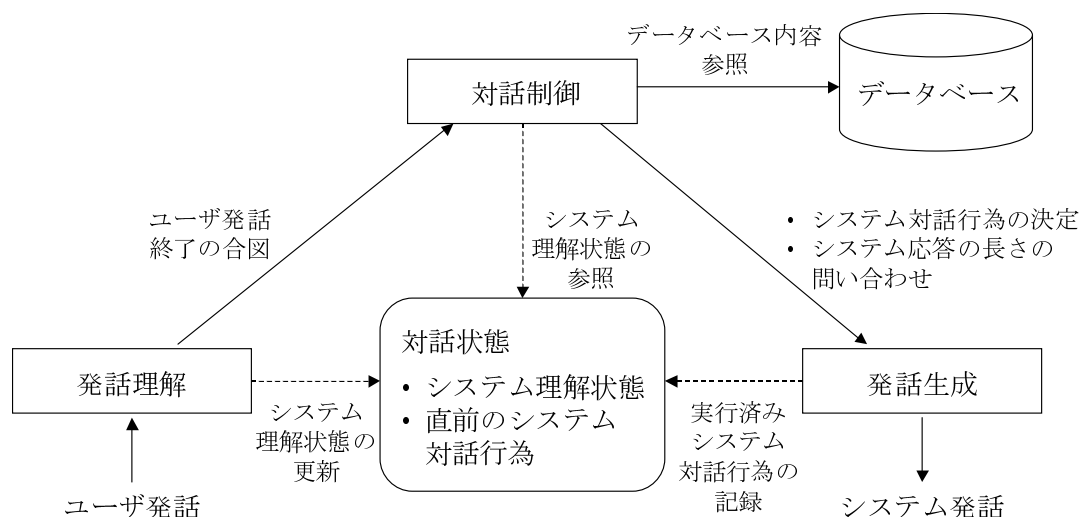


図 5.4: デュアルコスト法のシステム構成

発話理解部は、現在の対話状態の下でユーザ音声からユーザ発話内容を理解し、対話状態の中のシステム理解状態を更新する。システム理解状態の更新が終了したら、発話理解部は対話制御部に合図を送る。対話制御部は、対話の効率性の観点から、対話の各時点で最適なシステムの対話行為を決定する。発話生成部は、対話制御部の決定にしたがって、システム応答の言語表現を生成し、音声として出力する。実行済みのシステム対話行為は対話状態に書き込まれる。

対話制御部は、対話の各時点のシステム理解状態に基づいて、可能なすべての対話プランを列挙する。対話プランとはシステムがとるべき行為の列であり、対話の手順を表す。次に、各対話プランの対話コストを計算し、最小の対話コストをもつ対話プランにしたがって、実行すべきシステム対話行為を決定する。対話プランの対話コストとは、プラン通りに対話を進めたときに費やされる確認コストと情報伝達コストの和である。確認コストは属性認識精度に依存して計算される。属性認識精度とはシステムが属性の値を正しく認識できる確率である。情報伝達コストを計算するために、対話制御部は各対話プランにおいて予定されているシステム応答に含まれる自立語の数の期待値を発話生成部に対して問い合わせる。このように、対話制御部は、システム理解状態を随時更新する発話理解部と、予測されるシステム応答の長さの期待値を計算する発話生成部と連動しながら、コストの和が最小となるような対話を導くシステム対話行為を選択する。

- (Step1) 現時点のシステム理解状態において可能なユーザ問い合わせタイプを導き出し，ユーザ問い合わせタイプの確率分布を計算する．
- (Step2) ユーザ問い合わせタイプごとに，現時点のシステム理解状態の下で可能なすべての対話プランを生成し，各対話プランの対話コストを計算する．
- (Step3) 現時点で実行可能なシステム対話行為をすべて生成する．個々の対話行為に関して，ユーザ問い合わせタイプごとに，その対話行為を直ちに実行できるような対話プランの中で最小のコストをもつ対話プランを探索する．そのコストをシステム対話行為のコストと呼ぶ．
- (Step4) ユーザ問い合わせタイプの確率分布に基づいて，各システム対話行為のコストの期待値を計算し，コストの期待値が最小となるシステム対話行為を選択する．
- (Step5) (発話生成部が) 選択されたシステム対話行為を実行する．
- (Step6) (発話理解部が) 新たなユーザ発話を受け付け，システム理解状態を更新し，(Step1) に戻る．

図 5.5: デュアルコスト法の動作手順

5.3.2 動作手順

デュアルコスト法の動作手順を図 5.5 に示す．(Step1) においては，現時点のシステム理解状態で承認済みとなっている属性値が無効となるようなユーザ問い合わせタイプを排除することにより，可能なユーザ問い合わせタイプを導き出す．ユーザ問い合わせタイプの確率分布については 5.3.3 節で説明する．

(Step2) において，1 つの対話プランは，確認行為と情報要求行為の任意回の繰り返しと，それに続く一つの情報伝達行為から成る．確認行為と情報要求行為はなくてもよい．デュアルコスト法では，確定型の情報伝達行為を含んだ対話プランのみを考慮する．なお，確定型の情報伝達行為を含む対話プランを確定型の対話プランと呼び，試行型の情報伝達行為を含んだ対話プランを試行型の対話プランと呼ぶ．対話プランの生成法は 5.3.4 節で説明し，対話プランの対話コストの計

算法は5.3.5節で説明する．

(Step3)において，現時点で実行可能なシステム対話行為とは，承認済みでない属性値を確認するための確認行為，値が与えられていない属性の値を要求するための情報要求行為，現在のシステム理解状態に基づいて応答を生成する情報伝達行為である．システムの対話行為を直ちに実行できる対話プランとは，その行為を先頭にもつような対話プランである．

(Step4)では，対話の長さを最小にするという基準の下で最適なシステム対話行為が選ばれる．選ばれた対話行為が，確認行為や情報要求行為であるならば，ユーザ問い合わせ把握フェーズが継続し，情報伝達行為であるならば，(Step5)でシステム情報伝達フェーズに入ることになる．このようにして，デュアルコスト法は，ユーザ問い合わせ把握フェーズ(確認対話)からシステム情報伝達フェーズ(システム応答)への移行のタイミングを決定する．

5.3.3 ユーザ問い合わせタイプの確率分布

ここでは，ユーザ問い合わせタイプの確率分布を適切度 [89] という概念に基づいて定義する．

5.3.2節で説明したように，デュアルコスト法では，(Step1)において，現時点のシステム理解状態において可能なユーザ問い合わせタイプ $QUERY_j$ ($j = 1, \dots, m$) を導く．システム理解状態に含まれる各属性 $attr_i$ ($i = 1, \dots, n$) は値 v_i をとり，各属性の属性認識精度は r_i であるとする．ユーザ問い合わせタイプ $QUERY_j$ にとって有効な属性の個数を N_{QUERY_j} ，属性値 v_i が有効となるようなユーザ問い合わせタイプの個数を M_{v_i} とする．

ここで，ユーザ問い合わせタイプ $QUERY_j$ にとって属性値 v_i が有効であるとき真となる命題を $Valid_{ij}$ と書き，属性値 v_i が承認済みであるとき真となる命題を Ack_i と書くこととする．

このとき，ユーザ問い合わせタイプ $QUERY_j$ の現在のシステム理解状態に対する適切度 $Relevance(QUERY_j)$ を次のように定義する．

$$Relevance(QUERY_j) = \frac{1}{N_{QUERY_j}} \sum_{i=1}^n \frac{c_{ij}}{M_{v_i}} \quad (5.3)$$

where

$$c_{ij} = \begin{cases} r_j & Valid_{ij} \wedge \neg Ack_i \\ 1 & Valid_{ij} \wedge Ack_i \\ 0 & \neg Valid_{ij} \end{cases}$$

デュアルコスト法は、各ユーザ問い合わせタイプの現在のシステム理解状態に対する適切度を正規化したものをユーザ問い合わせタイプの確率分布として用いる。

5.3.4 対話プランの生成

5.3.2節で説明したように、(Step1)で、現在のシステム理解状態から可能なユーザ問い合わせタイプを導出した後、(Step2)では、ユーザ問い合わせタイプの各々について可能な対話プランを網羅的に生成する。対話プランは、確認行為と情報要求行為の任意回の繰り返しと、それに続く一つの情報伝達行為から成る。本章では簡単のため、情報要求行為として一時に唯一の属性の値を要求する情報要求行為のみを考える。

以下において、生成すべき対話プランを定義する。まず、ユーザ問い合わせタイプにとって有効な属性の集合を A_{valid} とする。有効な属性の集合 A_{valid} の中で現時点において値が与えられており、且つ、値が承認済みとなっている属性の集合を A_{ack} とする。現時点において値は与えられているが、承認済みでない属性の集合を A_{unack} 、現時点において値が与えられていない属性の集合を A_{empty} とする。次の式が成り立つ。

$$A_{valid} = A_{ack} + A_{unack} + A_{empty} \quad (5.4)$$

一つの対話プランの中で、確認行為を施す属性集合として、 A_{unack} の任意の部分集合 U_{unack} を選択し、情報要求行為を施す属性集合として、 A_{empty} の任意の部分

分集合 U_{empty} を選択する． U_{unack} と U_{empty} に含まれる属性は，プランを実施する中で確認行為と情報要求行為によって値が承認される．

$$U_{unack} \subseteq A_{unack} \quad (5.5)$$

$$U_{empty} \subseteq A_{empty} \quad (5.6)$$

U_{unack} を任意に分割し，それぞれの分割に対する確認行為を生成し， U_{empty} に含まれる属性の一つずつに対して情報要求行為を生成する．このようにして生成される確認行為と情報要求行為の任意の列を含むような対話プランが生成される．

次に，対話プランに含まれる情報伝達行為を生成するために，情報伝達行為の基礎となる属性集合を選択する必要がある．確定型の対話プランの場合，プランを実行する前に既に値が承認されている属性 A_{ack} と，プランを通して値が承認される U_{unack} と U_{empty} の総和を基礎とする確定型の情報伝達行為を生成する．

試行型の対話プランの場合，承認済みでない属性の値も正しいと仮定して，応答は生成される．そのため，確認対話が終わったとき，値は与えられているが，承認済みでない属性の集合 $A_{unack} - U_{unack}$ の任意の部分集合 V_{unack} を選択し， A_{ack} ， U_{unack} ， U_{empty} ， V_{unack} の総和を基礎とする試行型の情報伝達行為を生成する．

以上のような条件を満たす対話プランを網羅的に生成する．デュアルコスト法の場合，確定型の対話プランのみを生成する．なお，本章では属性を確認する順序によって確認コストは変わらないと仮定する．

たとえば，システム理解状態 S_1 においては，情報種別属性の値は警報であるが，承認済みではないので，可能な問い合わせタイプは，警報，天気，気温，降水確率の問い合わせである．このうち，警報問い合わせにおける確定型の対話プランの例としては，情報種別を確認してから，警報がどこにも発表されていないという応答を行うプラン $Plan_1$ と，場所と情報種別を確認してから，承認された場所には警報が発表されていないという応答を行うプラン $Plan_2$ がある．2つのプランは次のようにシステムの対話行為の列として表記される．

$$Plan_1 := [Act_1, Res_1]$$

$$Plan_2 := [Act_2, Act_1, Res_2]$$

ただし，

$Act_1 :=$ 情報種別属性についての情報要求行為

$Act_2 :=$ 場所属性についての確認行為

$Res_1 :=$ どこにも警報が発表されていないことを伝達するという
情報伝達行為

$Res_2 :=$ 承認された場所には警報が発表されていないことを伝達する
という情報伝達行為

警報問い合わせにおける試行型の対話プランの例としては，確認行為，情報要求行為なしで，情報種別属性の値が警報であると仮定して，どこにも警報が発表されていないことを伝達するプラン $Plan_3$ がある．

$Plan_3 := [Res_3]$

ただし，

$Res_3 :=$ 情報種別属性の値が警報であると仮定して，どこにも警報
が発表されていないことを伝達するという情報伝達行為

5.3.5 対話コストの計算

対話プランのコストの計算法について説明する．対話プランのコストを計算するためには，対話プランに含まれる確認行為，情報要求行為の確認コスト，確定型の情報伝達行為の情報伝達コストを計算し，その総和を求めればよい．

確認行為，情報要求行為の確認コストを計算するために，まず，一つの確認行為あるいは情報要求行為が完了するまでに要するシステム - ユーザ発話対の数の期待値について考える．一つのシステム - ユーザ発話対は，システムの確認発話 (例：「神奈川県ですか?」) あるいは情報要求への応答 (例：「いつですか?」) と，ユーザの訂正発話 (例：「香川県です」) あるいは情報伝達発話 (例：「今日です」) あるいは承認発話 (例：「はい」) から成る．

システムが各属性の値を正しく認識する確率を属性認識精度とよぶ．属性認識精度は前もって与えられていることを前提とする．また，属性の集合が与えられ

るとき，その集合に含まれるすべての属性の値を一度に正しく認識する確率を属性集合の認識精度とび，各属性の属性認識精度の積として計算できるものとする．また，システムはユーザの承認発話を常に正しく認識できるものと仮定する．

属性の集合が与えられるとき，その集合に含まれるすべての属性の値が承認済みとなるための確認行為について考える．システムとユーザの対話は，システムがすべての属性値を一度に提示することにより確認発話を行い，システムが提示した属性の値が一つでも誤っているなら，ユーザはすべての属性値を提示することにより訂正発話を行うという対話を繰り返していき，システムの提示する属性値がすべて正しければ，ユーザは承認発話を行い，そこで確認行為が完了するという動作系列であると仮定する．属性集合の認識精度を r とするとき，確認行為が完了するまでに要するシステム - ユーザ発話対の数の期待値は次の式で与えられる [89] ．

$$Pair_c = \sum_{i=1}^{\infty} ir(1-r)^{i-1} = \frac{1}{r} \quad (5.7)$$

情報要求行為は，属性の値が与えられていないときに，最初に情報要求発話を 1 回行い，その後は確認行為と同じ動作系列をとる．したがって，情報要求行為が完了するまでに要するシステム - ユーザ発話対数の期待値は次の式で与えられる．

$$Pair_s = 1 + Pair_c = 1 + \frac{1}{r} \quad (5.8)$$

確認行為，情報要求行為の確認コストは，行為を遂行するために要する各システム - ユーザ発話対に含まれる自立語の数の総和である．ここでは，確認発話，情報要求発話，訂正発話においては，一つの属性が一つの名詞として実現され，その名詞のみが発話に含まれる自立語であると仮定する．また，承認発話は，1 個の肯定的な応答表現 (例: 「はい」) によって実現されると仮定する．

まず，確認行為の確認コストについて考える．システム確認発話，ユーザ訂正発話においては，確認の対象となっているすべての属性値が言及されると仮定していることに注意する．確認すべき属性の数を m 個とすると，最後の 1 回を除いて，システム - ユーザ発話対は， m 個の自立語を含むシステム確認発話と， m 個

の自立語を含むユーザ訂正発話から成る．最後 1 回のシステム - ユーザ発話対は， m 個の自立語を含むシステム確認発話と，1 個の自立語を含むユーザ承認発話から成る．したがって，一つの確認行為が完了するまでにやり取りされる自立語の数の期待値は次の式で表される．

$$2m(Pair_c - 1) + m + 1 = \frac{2m}{r} - m + 1 \quad (5.9)$$

同様に考えて，一つの情報要求行為の確認コストは次の式で表される．

$$2m(Pair_s - 1) + m + 1 = \frac{2m}{r} + m + 1 \quad (5.10)$$

対話プランの確認コストは，確認行為と情報要求行為の確認コストの計算式 (5.9), (5.10) を使って，対話プランに含まれるすべての確認行為と情報要求行為の確認コストの和として計算できる．

次に，確定型の情報伝達行為に対する情報伝達コストの算出について説明する．情報伝達コストは，情報伝達行為を実施するために生成されるシステム応答に含まれる自立語の数の期待値である．たとえば，情報伝達行為が「警報はどこにも発表されていない」というシステム応答として実行されるなら，「警報」，「どこにも」，「発表されていない」という 3 つの自立語が含まれるので，情報伝達コストは 3 となる．対話プランによっては確認行為や情報要求行為によって確定される属性の値が複数通りあり，対話を実際に行ってみないことには値が決まらず，情報伝達行為として生成されるシステム応答も一意に決まらないという場合がある．こういった場合には，属性がとりうる値の生起確率は等確率であると仮定した上で，すべての値の組合せを考慮したときのシステム応答に含まれる自立語の数の期待値が情報伝達コストであると定義する．

5.4 試行型デュアルコスト法

試行型デュアルコスト法は，5.3 節で説明したデュアルコスト法と同じ手順で動作する．デュアルコスト法と異なる点は，試行型デュアルコスト法は，承認済み

でない問い合わせ内容も正しいと仮定して応答を生成するという試行型の対話プランも考慮するということのみである．ここでは，試行型の対話プランのコストの計算法について説明する．

試行型の対話プランは，確認行為あるいは情報要求行為の任意回の繰り返しと，それに続く試行型の情報伝達行為から成る．確認行為，情報要求行為の確認コストは，5.3 節で示した計算式により計算できる．以下においては，試行型の情報伝達行為に関わるコストの計算法について述べる．

試行型の情報伝達行為は，承認済みでない，いくつかの属性の値が正しいという仮定の下に，試行型のシステム応答を実行することから始まる．仮定が真ならば，システムはユーザが必要とする情報を伝達することができ，試行型のシステム応答は成功する．仮定が偽であるなら，試行型のシステム応答は失敗する．試行型のシステム応答が成功する確率をシステム応答の成功確率と呼ぶ．

試行型のシステム応答は，応答が失敗したときにユーザが気づくように，ユーザ問い合わせ内容に明示的に言及するような表層形式として生成されると仮定する．さらに，ユーザは，システム応答の失敗に気づいたときには，必要とする情報を手に入れるまで，同じ問い合わせを繰り返すと仮定する．

ここで，ユーザがタイプ τ の問い合わせ Q を行う状況について考える．システムが，属性集合 A_0 に含まれる属性の値が正しいと仮定して，試行型のシステム応答 R_0 を生成したと想定する．試行型の応答 R_0 に含まれる自立語を I_0 とする． I_0 は，応答 R_0 が成功するか否かにかかわらず，必ず費やされることになる情報伝達コストである．応答 R_0 の成功確率を q_0 とする．成功確率 q_0 は属性集合 A_0 に含まれるすべての属性の認識精度の積であると近似する．応答 R_0 が確率 q_0 で成功した場合，対話はそこで終了する．応答 R_0 が確率 $1 - q_0$ で失敗した場合，タイプ τ の問い合わせ Q に対処する対話の列 $D_1, D_2, \dots$ が，ユーザが必要な情報を手に入れるまで続くことになる．個々の対話 D_i は，試行型デュアルコスト法がタイプ τ の問い合わせに対処する対話である．

ここで，試行型デュアルコスト法がタイプ τ の問い合わせに対処するための 1 回の対話において，確認対話でやり取りされる自立語の数の平均を $\overline{C}_\tau$ とし，システム応答中の自立語の数の平均を $\overline{I}_\tau$ とする．別の言い方をすれば， $\overline{C}_\tau$ は，タイプ τ の問い合わせに対処するための 1 回の対話における確認コストの平均であり， $\overline{I}_\tau$

は情報伝達コストの平均である．また，タイプ τ の問い合わせに対処するときの応答成功確率の平均を $\bar{q}_\tau$ とする．

以上の考察より，応答 R_0 が成功するか否かに関わらず，情報伝達コスト I_0 がかかり，応答 R_0 が確率 $1 - q_0$ で失敗した場合には， $\bar{C}_\tau + \bar{I}_\tau$ 個の自立語を含む対話が， $1/\bar{q}_\tau$ 回続くことになる．結果として，試行型の応答 R_0 から始まる試行型の情報伝達行為に関わるコストの総和は次の式で与えられる．

$$I_0 + \frac{1 - q_0}{\bar{q}_\tau} (\bar{C}_\tau + \bar{I}_\tau) \quad (5.11)$$

5.5 シミュレーション対話実験による評価

5.5.1 シミュレーション対話実験

システムと模擬ユーザとの間のシミュレーション対話実験によって，デュアルコスト法，試行型デュアルコスト法の評価を行った．模擬ユーザとは，実ユーザの振舞いをシミュレートしながらシステムと対話するエージェントである．各対話の初期時点において，模擬ユーザはシステムに対する問い合わせ内容を保持している．ユーザはシステムに問い合わせ内容を伝え，システムは確認対話を通して問い合わせ内容を把握していく．問い合わせ内容の把握が完了すると，システムはユーザ問い合わせ内容に応じてユーザが必要とする情報を伝達するためにシステム応答を生成する．システム応答が模擬ユーザが必要とする情報を伝えていないなら，模擬ユーザは再度同じ問い合わせ内容で対話を繰り返す．対話は模擬ユーザが必要とする情報が伝達されるまで継続する．

システムと模擬ユーザは，音声で会話するのではなく，発話内容を表現した属性と値の対のリストをやり取りすることによって会話する．ユーザの発話内容をシステムに送るときには，属性認識精度に応じて属性値に誤りが含まれるように，システムの音声認識誤りをシミュレートした．

実験で用いたタスクは，5.2 節で述べた天気情報案内タスクである．場所は約 700 の都市，日は今日か明日の 2 通りである．警報種別としては，洪水，大雨など 10 個の種別がある．システムは，データベースの中に，各都市の今日，明日の天気，

- (C-1) システム理解状態から可能なユーザ問い合わせタイプを導き出す．(C-2) に移行する
- (C-2) ユーザ問い合わせタイプが一意に決まっており，その問い合わせタイプの属性値がすべて承認済みなら，システム情報伝達フェーズに移行する．さもなければ，(C-3) に移行する．
- (C-3) 一括確認か個別確認かに応じて，次のいずれかを実行する．
一括確認の場合：システム理解状態において，既に値が与えられている属性があるなら，そういった属性の値をできるだけ多く一度に確認するための確認行為を選択し，(C-5) へ移行する．さもなければ，(C-4) へ移行する．
個別確認の場合：システム理解状態において，既に値が与えられている属性があるなら，それらの属性の一つを確認するための確認行為を選択し，(C-5) へ移行する．さもなければ，(C-4) へ移行する．
- (C-4) できるだけ多くのユーザ問い合わせタイプで有効となる属性を優先するように，値が与えられていない属性の一つを選択し，その属性のための情報要求行為を選択する．(C-5) に移行する．
- (C-5) 選択されたシステムの対話行為を実行し，ユーザからの応答を待って，システム理解状態を更新する．(C-1) へ戻る．

図 5.6: 一括確認法と個別確認法の動作手順

最高気温と最低気温，6 時間ごとの降水確率のデータを保持している．また，警報については，現在どこにも警報は発表されていないということをデータとして保持している．このデータベースの内容の場合，警報の問い合わせに関しては，情報種別だけを確認することが，最適な対話プランとなる．天気，気温，降水確率の問い合わせに関しては，たいていの場合，場所，日，情報種別の属性をすべてを確認することが最適な対話プランとなる．ただし，天気問い合わせに関しては，日属性の確認を行わずに，今日と明日の両日の気象情報を伝達することが，最適なプランとなる場合もある．これは，属性認識精度によっては，日属性の確認を確認を行うよりも，確認を省いて両日の天気を伝達した方が，対話コストが小さ

- (N-1) システム理解状態から可能なユーザ問い合わせタイプを導き出す．(N-2) に移行する．
- (N-2) システム理解状態における属性値がすべて承認されているものと仮定したときに，ユーザ問い合わせタイプが一意に決まるなら，システム情報伝達フェーズに移行する．さもなければ，(N-3) に移行する．
- (N-3) できるだけ多くのユーザ問い合わせタイプで有効となる属性を優先するように，値が与えられていない属性を一つ選択し，その属性のための情報要求行為を選択する．(N-4) に移行する．
- (N-4) 選択されたシステムの対話行為を実行し，ユーザからの応答を待って，システム理解状態を更新する．(N-1) へ戻る．

図 5.7: 無確認法の動作手順

くなる場合があるからである．

デュアルコスト法，試行型デュアルコスト法と比較するために，音声認識精度とデータベース内容に依存しない対話制御法として，一括確認法，個別確認法，無確認法と呼ぶ対話制御方法を用いた．一括確認法と個別確認法は確定型の対話戦略をとる．一括確認法は，できるだけ多くの属性を一度に確認しようとする方法であり，個別確認法は，属性を一つずつ確認する方法である．無確認法は確認を行わない方法であり，試行型の対話戦略をとる．一括確認法と個別確認法の動作手順を図 5.6 に示し，無確認法の動作手順を図 5.7 に示す．

模擬ユーザの振る舞いは図 5.8 に示すルールにしたがって決定される．(U-4) は，ユーザ問い合わせタイプが天気問い合わせであるにもかかわらず，システムが警報種別属性の値について，確認発話や情報要求発話を行ってくるような場合に相当する．そういった場合，模擬ユーザはシステムの確認発話あるいは情報要求発話に応えることができないことを伝えるための発話(拒否発話)を行う．システムは，模擬ユーザの拒否発話を受け取ると，現在の行為をあきらめ，別のシステム発話を選択し実行する．

デュアルコスト法，試行型デュアルコスト法，一括確認法，個別確認法，無確

- (U-1) 対話の開始時点では，問い合わせ内容の一部をシステムに伝える．
- (U-2) システムの確認発話に対して，訂正発話か承認発話を行う．訂正発話は，システムの確認発話に含まれるすべての属性の値を伝達することによって行う．
- (U-3) システムの情報要求発話に対して，属性の値を伝達するための発話を行う．
- (U-4) システムがユーザ問い合わせタイプにとって無効な属性の値について，確認発話あるいは情報要求発話を行ってきたならば，システム発話を拒否する．

図 5.8: 模擬ユーザの振る舞いを決定するルール

認法のそれぞれにおいて，システムの確認発話や情報要求発話が模擬ユーザから拒否された場合に，別のシステム発話を選択する際の基準を説明する．デュアルコスト法，試行型デュアルコスト法では，次にコストが小さい行為を優先して選択する．一括確認法では，図 5.6 において，(C-3) で選んだ確認行為が拒否されたなら，まだ確認を試みていない属性の組合せのうち，できるだけ多くの属性の値を一度に確認する確認行為を優先して選択する．(C-4) で選んだ情報要求行為が拒否されたなら，まだ情報要求を試みていない属性のうち，できるだけ多くのユーザ問い合わせタイプで有効となる属性を優先して選び，その属性のための情報要求行為を選択する．個別確認法では，図 5.6 において，(C-3) で選んだ確認行為が拒否されたなら，値が与えられている別の属性を任意に選び，その属性についての確認行為を選択する．(C-4) で選んだ情報要求行為が拒否された場合は，一括確認法と同様の方法で別の情報要求行為を選択する．無確認法では，図 5.7 において，(N-3) で選んだ情報要求行為が拒否されたなら，一括確認法，個別確認法と同様の方法で別の情報要求行為を選択する．

シミュレーション対話実験では，4 つのユーザ問い合わせタイプごとにユーザ問い合わせ内容を無作為に生成した．各属性の属性認識精度を等しく 0.75 から 1.0 まで 0.5 刻みで変化させた．各属性認識精度において 2000 回のシミュレーション対話が実施された．試行型デュアルコスト法を動作させるためには，式 (5.11) に示したように，失敗したシステム応答に後続する対話について，システム応答成功確率の平均 $\overline{q}_\tau$ ，対話の長さの平均値， $\overline{C}_\tau + \overline{I}_\tau$ の値が必要となる．これらの値と

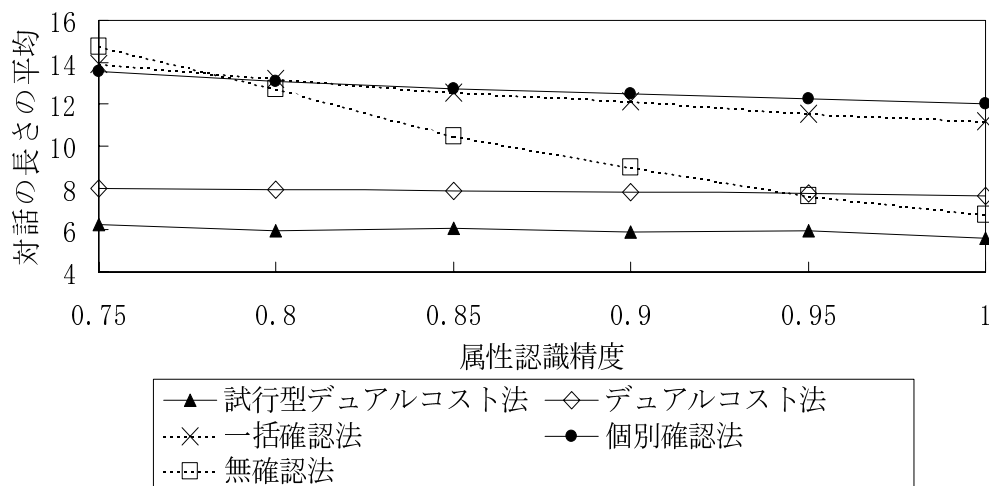


図 5.9: 警報問い合わせの場合における属性認識精度に応じた対話の長さの平均

しては、試行型デュアルコスト法のシミュレーション対話実験を事前に何度か繰り返し、収束した値を用いた。

各対話制御法の性能はユーザが必要とする情報が伝達されるまでの対話の長さの平均によって評価した。対話の長さは、5.3.5 節で述べた基準に加えて、模擬ユーザの振舞 (U-4) におけるユーザ拒否発話は簡潔な否定的表現 (例:「いいえ」、「分かりません」) として実現されると仮定した上で、拒否発話の長さは 1 であるという基準にしたがって計算した。

5.5.2 結果と考察

図 5.9, 図 5.10, 図 5.11, 図 5.12, 図 5.13 にシミュレーション対話実験の結果を示す。それぞれのグラフは、属性認識精度に応じた対話の長さの平均の推移を示している。図 5.9 は問い合わせタイプが警報問い合わせの場合、図 5.10 は問い合わせタイプが天気問い合わせの場合、図 5.11 は問い合わせタイプが気温問い合わせの場合、図 5.12 は問い合わせタイプが降水確率問い合わせの場合、図 5.13 は 4 つの問い合わせタイプを無作為に発生させた場合の結果である。

問い合わせタイプによってデータベース内容の状態が異なるので、デュアルコスト法、試行型デュアルコスト法の効果も異なる。警報問い合わせの場合は、場

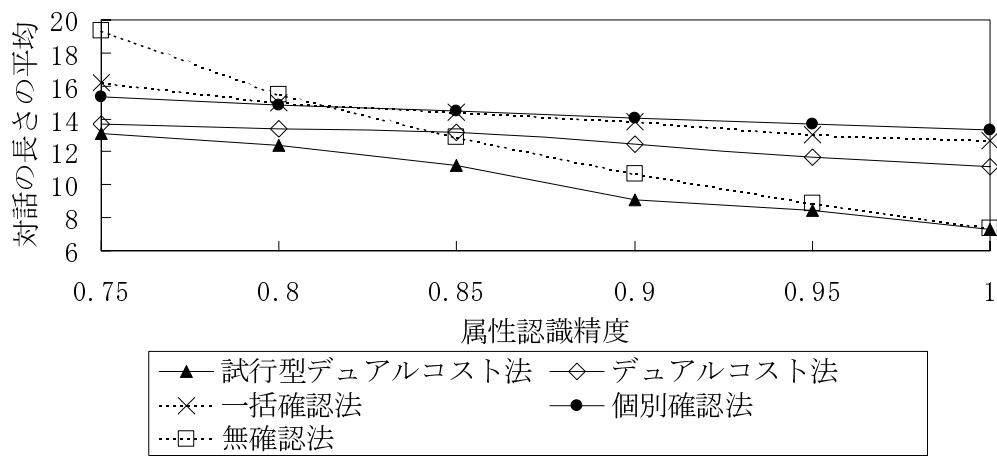


図 5.10: 天気問い合わせの場合における属性認識精度に応じた対話の長さの平均

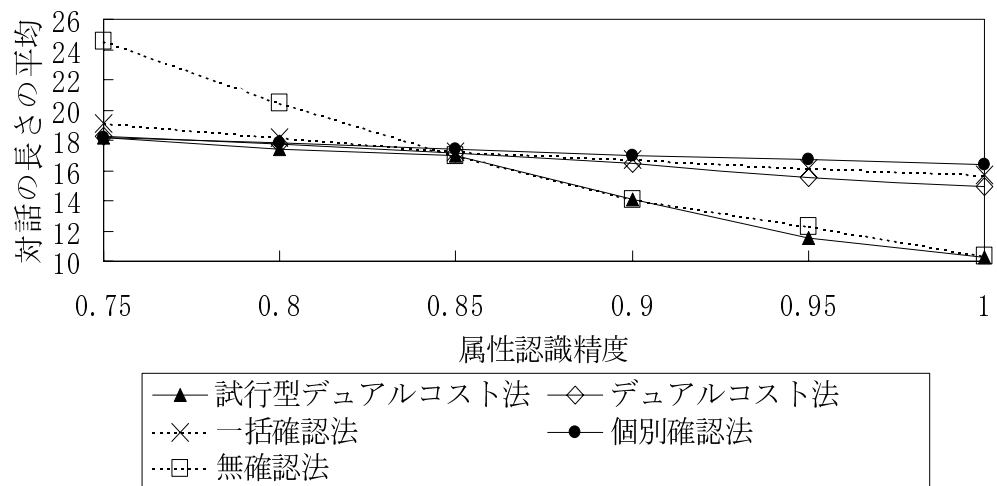


図 5.11: 気温問い合わせの場合における属性認識精度に応じた対話の長さの平均

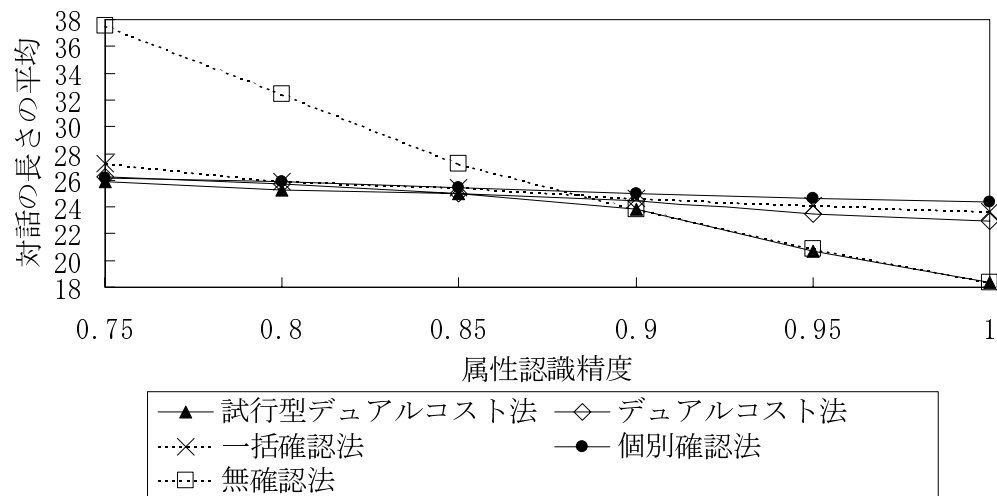


図 5.12: 降水確率問い合わせの場合における属性認識精度に応じた対話の長さの平均

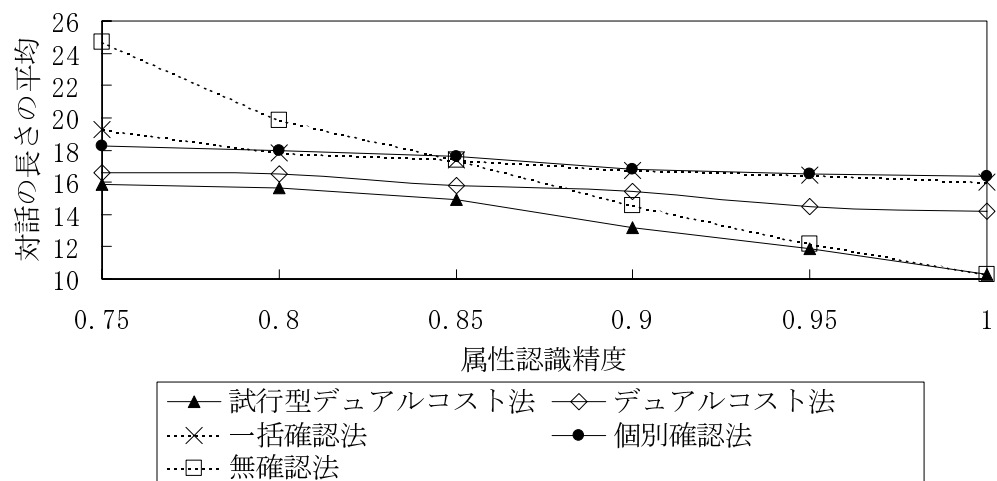


図 5.13: 各問い合わせタイプを無作為に発生させた場合における属性認識精度に応じた対話の長さの平均

所属性の確認を回避できるという点で、(試行型)デュアルコスト法の効果が最も発揮されやすい場合である。気温や降水確率の問い合わせの場合は、(試行型)デュアルコスト法であっても、場所属性、日属性、情報種別属性のすべてを確認しなければならない場合がほとんどであり、(試行型)デュアルコスト法の効果を発揮することは困難である。天気問い合わせの場合は、(試行型)デュアルコスト法は、日属性の値の確認を省いて、今日と明日の天気を伝達するということがありえる。このように、様々なデータベースの場合を取り上げることにより、(試行型)デュアルコスト法にとって有利な状況では、(試行型)デュアルコスト法が実際に効果を上げることができ、そうでない状況であっても、従来法に比べて対話の効率を低下させないことを実証することができる。

まず、確定型の対話戦略のみを考慮するデュアルコスト法、一括確認法、個別確認法を比較する。図 5.9 から分かるように、警報問い合わせの場合、デュアルコスト法は、一括確認法、個別確認法と比較して、より短い対話で情報を伝達できた。警報問い合わせの場合には、警報がどこにも発表されていないというデータベースの内容であることを利用して、デュアルコスト法は、一括確認法、個別確認法では避けることができない不必要な確認を回避することによって、効率的に対話を実施できたことが分かる。この実験においては、システムは模擬ユーザ発話を正しく認識できるとは限らず、模擬ユーザが警報の問い合わせを行ったとしても、システムは天気、気温、降水確率の問い合わせであると誤認識する場合がある。デュアルコスト法は、そういった誤認識が起きる場合であっても、対話全体の効率性を向上させるように対話を制御できたことが分かる。

図 5.10 から分かるように、天気問い合わせの場合は、警報問い合わせの場合ほどではないが、デュアルコスト法は、一括確認、個別確認法よりも短い対話で情報を伝達できたことが分かる。これは、日属性の値を確認しないで、今日と明日の両日の天気を伝達するという戦略が有効に働くためである。

図 5.11、図 5.12 から分かるように、気温、降水確率の問い合わせの場合には、デュアルコスト法による対話の長さは、一括確認法とほぼ同じである。これは、気温問い合わせ、降水確率の場合には、デュアルコスト法であろうと、場所属性、日属性、情報種別属性をすべて確認しなければならない場合がほとんどであるからである。しかし、データベースの内容によらず問い合わせ内容を逐一確認しなけ

ればならないような場合であっても，デュアルコスト法が，一括確認法に比較して対話の効率を低下させることはないことが分かる．

以上の考察から，確定型の対話戦略のみを考慮する場合，デュアルコスト法は，効果を発揮しやすいデータベース内容の場合には一括確認法，個別確認法よりも効率的な対話を実施することができ，効果を発揮することが困難なデータベース内容であっても，一括確認法，個別確認法よりも対話の効率を低下させることはないことが分かる．

次に，試行型の対話戦略を考慮する試行型デュアルコスト法，無確認法を加えて，5つの対話制御法を比較する．図 5.9，図 5.10，図 5.11，図 5.12 から分かるように，すべての問い合わせタイプについて，属性認識精度が高くなると，無確認法がデュアルコスト法よりも短い対話を実現できている．警報問い合わせの場合には，デュアルコスト法にとってもっとも有利なデータベースの状況であるので，デュアルコスト法が無確認法に劣るのは認識精度が1に近い場合のみであるが，デュアルコスト法が有利に働くようなデータベースの内容でない場合には，認識精度が0.85を超えると，デュアルコスト法よりも無確認法の方が短い対話を実現している．

その一方で，デュアルコスト法を改良した試行型デュアルコスト法は，すべての問い合わせタイプについて，属性認識精度によらず，他の4つの方法よりも短い対話でユーザが必要とする情報を伝達できていることが分かる．試行型デュアルコスト法は，属性認識精度が低い場合には，デュアルコスト法と似た性能を示し，認識精度が高くなると，無確認法と似た性能を示している．これは，試行型デュアルコスト法が，属性認識精度が低いときには確定型の対話戦略を採用し，認識精度が高い場合には試行型の対話戦略をとるように働くためである．無確認法は，認識精度が低いと応答の失敗が多くなり，デュアルコスト法は，認識精度が高いと十分すぎるほど確認を行って，かえって対話の効率を落とすが，試行型デュアルコスト法は，認識精度に応じて適切な対話戦略を選択することにより，すべての認識精度で最も短い対話でユーザに必要な情報を伝達することが可能であることが分かる．図 5.13 から分かるとおり，4つの問い合わせタイプを無作為な順序で与えたとしても，同様の結果を得る．

この実験においては，4つの属性から成るタスクを用いたが，属性の数はデュア

ルコスト法，試行型デュアルコスト法にとっては本質的なことではない．属性の数が多くなっても，ここで用いた天気情報案内タスクのように，対話時点のデータベースの内容に依存して対話を制御しないことには，短い対話でユーザの必要とする情報を伝達することができないようなタスクであれば，デュアルコスト法，試行型デュアルコスト法は効果を発揮する．

5.6 被験者対話実験による評価

デュアルコスト法，試行型デュアルコスト法の効果を被験者対話実験によって評価した．評価用のシステムは，日本語 SpeechBuilder を利用して構築した [32]．発話理解部，発話生成部については，日本語 SpeechBuilder によって生成されるものを利用した．対話制御部は，デュアルコスト法，試行型デュアルコスト法をはじめ，5.5 節で述べた一括確認法，個別確認法，無確認法を切り替えて動作させることが可能なものを作成した．対話制御部と他のモジュールの連動は，SpeechBuilder が提供する CGI プロトコルを利用して実装した．

評価システムのタスクは，5.2 節で述べた天気情報案内タスクである．ユーザ問い合わせのタイプは，天気，気温，降水確率，気温の問い合わせの 4 種類であり，システム理解状態は，場所，日，警報種別，情報種別の 4 つの属性によって表現される．システムのデータベースには，日本全国の都道府県とすべての市について，今日，明日ごとの，天気，最高・最低気温，6 時間ごとの降水確率と，現在発表されている警報のデータが記録されている．実験においては，どこにも警報が発表されていないというデータベースの状態で行った．システムの認識語彙は約 800 語である．

以下において，属性認識精度と式 (5.11) における定数 $\overline{q_r}$, $\overline{C_r} + \overline{I_r}$ の値を獲得するために行った予備実験について説明し，次に，それらの獲得された値を用いて行った評価実験について説明する．評価実験の結果について考察し，デュアルコスト法，試行型デュアルコスト法の有効性を論じる．

5.6.1 予備実験

各属性の認識精度を獲得するための予備実験を行った。この予備実験では、3名の被験者の各々に対してあらかじめ決められた20個の問い合わせが与えられ、合計60対話を実施された。問い合わせタイプは無作為に選ばれた。各被験者は与えられた問い合わせを順にシステムに対して行った。

システムはデュアルコスト法によって動作させた。デュアルコスト法を動作させるためには、属性の認識精度を与える必要があるが、予備実験の段階では正しい認識精度は分からないので、各対話において0.7から1.0までの値を無作為に選んで、すべての属性の認識精度をその値に設定し、デュアルコスト法を動作させた。

属性の認識精度の獲得法について説明する。各対話において、対話終了時のシステム理解状態で承認済みの属性値は正しいと仮定した。対話途中で認識された属性値が、それらの承認済みの値と矛盾しないなら、正しく認識されたと見なし、さもなければ誤って認識されたものとみなすことによって、属性の認識精度を獲得した。結果として、場所属性、日属性、警報種別属性、情報種別属性の認識精度は、それぞれ、0.62, 0.96, 0.93, 0.94であった。

次に、式(5.11)における定数 $\overline{q_\tau}$, $\overline{C_\tau} + \overline{I_\tau}$ の値を獲得するために、別の予備実験を行った。この予備実験でも、第一の予備実験と同じように、3名の被験者の各々に対して、あらかじめ決められた20個の問い合わせが与えられ、合計で60対話が順に実施された。問い合わせタイプは無作為に選ばれた。システムの対話制御部は、試行型デュアルコスト法によって動作させた。属性の認識精度は、第一の予備実験で獲得したものをを用いた。予備実験開始時には、 $\overline{q_\tau}$ の値を1に設定し、 $\overline{C_\tau} + \overline{I_\tau}$ の値は、第一の予備実験において、各対話ごとにシステムと被験者がやり取りした自立語の数の平均値に設定した。各対話が終了するごとに、これらの定数の値をその時点までに観察された対話ごとの値の平均値に更新した。60対話が終了した時点で得られた値を評価実験で用いた。

5.6.2 評価実験

デュアルコスト法、試行型デュアルコスト法の効果を実証するために評価実験を行った。2つの方法の効果の実証するために、一括確認法、個別確認法、無確認

法と比較した。デュアルコスト法，一括確認法，個別確認法は，確定型の対話戦略のみを考慮する方法であり，試行型デュアルコスト法と無確認法は，試行型の対話戦略を考慮する方法である。また，一括確認法と個別確認法は，音声認識精度，データベース内容にかかわらず，すべての属性を確認する方法である。

実験では，15名の被験者の各々に対して，あらかじめ決められた20個の問い合わせが与えられ，合計300対話が実施された。各問い合わせについて，システムが用いる対話制御法はあらかじめ決められている。ユーザはシステムがどの対話制御法で動作しているのかは知らない。システムは被験者に与えられた問い合わせが何かは知らない。問い合わせごとに適用される対話制御法は，4つの問い合わせタイプの各々に対して5つの対話制御法が公平に分配されるように選ばれた。

5.6.3 結果と考察

評価実験の結果を図5.14と表5.2に示す。図5.14と表5.2は同一の結果を示している。各問い合わせタイプごとに，試行型デュアルコスト法 (TDC)，デュアルコスト法 (DC)，一括確認法 (LS)，個別確認法 (P)，無確認法 (NC) で動作させたシステムと被験者が対話したときの対話の長さの平均 (被験者に必要な情報が伝達されるまでにやり取りされた自立語の数の平均) が示されている。対話の長さの平均値の差が統計的に有意であるかどうかを確かめるために，ノンパラメトリックな分散分析 (ANOVA) の一つである Kruskal-Wallis 法と，ノンパラメトリックな多重比較法の一つである Dunn 法を用いた ($p < 0.01$)。ノンパラメトリックな方法を用いたのは，対話の長さが正規分布に従うことを仮定できないためである。

まず，確定型の対話戦略のみを考慮する対話制御法として，デュアルコスト法，一括確認法，個別確認法の3つの方法を比較した。警報，天気問い合わせの場合，分散分析の結果，対話の長さの平均に有意な差があり，多重比較の結果，デュアルコスト法は他の2つの方法に比較して，対話の長さの平均が有意に小さいことが分かった。気温と降水確率の問い合わせに関しては，対話の長さの平均に有意な差は認められなかった。

この結果について考察する。まず，警報問い合わせの場合，デュアルコスト法は，どこにも警報が発表されていないというデータベース内容を利用して，不必

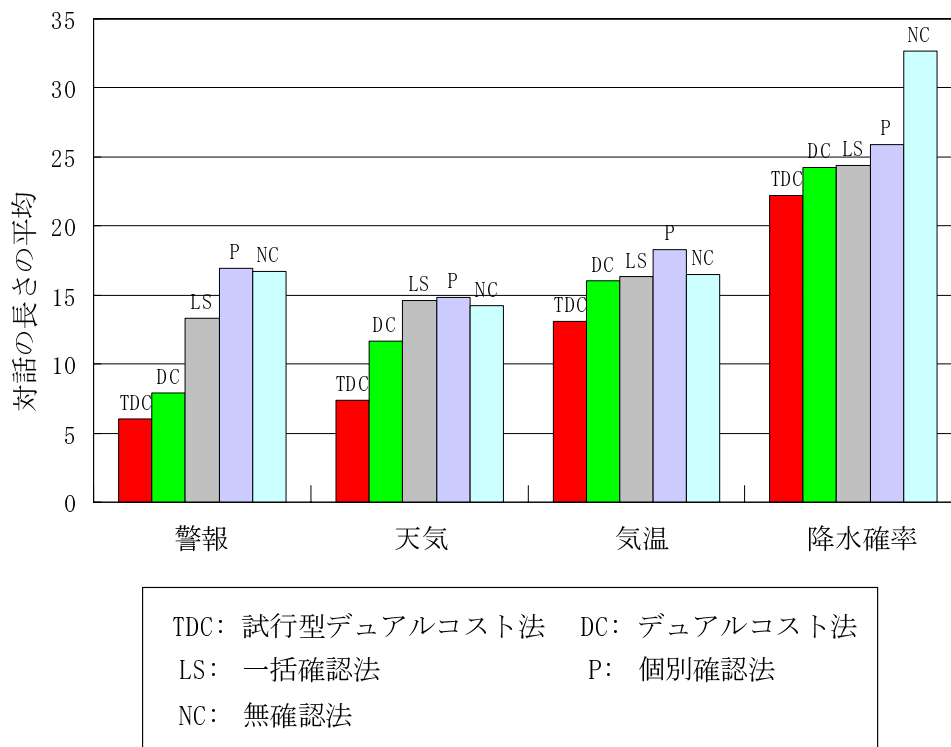


図 5.14: 被験者対話実験における各対話制御法による対話の長さの平均

要な確認を避けることができる。したがって、デュアルコスト法は、他の2つの方法よりも短い対話で必要な情報をユーザに伝達することができた。天気問い合わせの場合、デュアルコスト法は、日属性の値について確認を避け、今日と明日の天気を同時に伝達することを選ぶので、他の2つの方法よりも短い対話で必要な情報を伝達することができた。しかし、気温や降水確率の問い合わせの場合、天気の問い合わせに比較して、システム応答に含まれる自立語の数が多いため、デュアルコスト法であっても、日属性の確認を省くことはできない。したがって、気温、降水確率の問い合わせの場合は、デュアルコスト法であっても他の2つの方法よりも短い対話を実現することはできない。ここで重要なことは、そのような場合においても、デュアルコスト法の性能は他の2つの方法よりも悪くなることはないということである。

次に、試行型の対話戦略を考慮する試行型デュアルコスト法、無確認法を含め

て、5つの方法すべてについて比較を行った。警報、天気、気温の問い合わせの場合、分散分析の結果、対話の長さの平均には有意な差があり、多重比較の結果、試行型デュアルコスト法は他の4つの方法に比較して、対話の長さの平均が有意に小さいことが分かった。気温の問い合わせの場合、分散分析の結果、対話の長さの平均には有意な差があり、多重比較の結果、試行型デュアルコスト法は、無確認法を除く他の3つの方法に比較して、対話の長さの平均が有意に小さいことが分かった。試行型デュアルコスト法と無確認法との間には、対話の長さの平均に有意な差はなかった。降水確率の問い合わせの場合、分散分析の結果、5つの方法の間に対話の長さの平均に有意な差は認められなかった。

この結果について考察する。降水確率の問い合わせの場合、システムは6時間ごとの降水確率を伝達しようとするので、システム応答が非常に長くなる。そのため、試行型の応答が失敗したときのリスクが高くなるので、試行型デュアルコスト法は試行型の応答を選ぶことが少ない。したがって、降水確率の問い合わせの場合は、試行型デュアルコスト法であっても他の方法より短い対話でユーザに必要な情報を伝達できるわけではない。警報、天気、気温の問い合わせの場合は、試行型デュアルコスト法は試行型の応答を選択することによって短い対話を実現している。

表 5.2: 被験者対話実験における各対話制御法による対話の長さの平均

問い合わせタイプ	提案法		従来法		
	TDC	DC	LS	P	NC
警報	6.0	7.9	13.3	16.9	16.7
天気	7.4	11.7	14.6	14.8	14.2
気温	13.1	16.0	16.3	18.3	16.5
降水確率	22.2	24.2	24.4	25.9	32.7

TDC: 試行型デュアルコスト法 DC: デュアルコスト法
 LS: 一括確認法 P: 個別確認法
 NC: 無確認法

以上をまとめると、確定型の対話戦略だけを考慮するとき、デュアルコスト法は、データベース依存な対話を実施できる場合には従来法よりも短い対話でユーザに情報を伝達することが可能であり、そうでない場合であっても、従来法より対話の効率が低下することはない。試行型の対話戦略も考慮するとき、試行型デュアルコスト法は、試行型の応答のリスクが確認を省くことによる利得に見合う場合には、他の方法よりも短い対話で必要な情報を伝達することができ、そうでない場合であっても、他の方法に比べて対話の効率が低下することはない。

5.7 おわりに

本章では、音声認識精度とシステムのデータベースの内容に基づいて、できるだけ短い対話でユーザが必要とする情報を伝達することを可能とする対話制御法として、デュアルコスト法、試行型デュアルコストを提案した。デュアルコスト法は、ユーザの承認発話により承認済みとなった問い合わせ内容のみが正しいとして応答を生成する方法であり、試行型デュアルコスト法は、デュアルコスト法を発展させ、承認済みでない問い合わせ内容も正しいと仮定した上で応答を生成する方法である。両提案法は、対話の長さを確認コストと情報伝達コストの和として推定し、コストの和を最小にするように対話の各時点におけるシステムの対話行為を選択することにより、効率的な対話を実現する。対話制御部は、対話の各時点において、システム理解状態を随時更新する発話理解部と、予測されるシステム応答の長さを計算する発話生成部と連動しながら、コストの和が最小となるような対話を導くシステムの対話行為を選択する。結果として、従来法では回避できなかった不必要な確認を回避し、効率的な対話を実現することが可能となる。シミュレーション対話実験と被験者実験により、提案法が従来法よりも短い対話でユーザに必要な情報を伝達できることを示した。

- デュアルコスト法、試行型デュアルコスト法ともに、データベース内容に依存して、従来法が避けることはできない不必要な確認を回避することができる。
- シミュレーション対話実験と被験者実験の結果により、デュアルコスト法は、

音声認識精度によらず，確定型の対話戦略のみを考慮する一括確認法，個別確認法といった従来法よりも短い対話で必要な情報をユーザに伝達することができることが示された．

- シミュレーション対話実験と被験者実験の結果により，試行型デュアルコスト法は，音声認識精度によらず，デュアルコスト法，無確認法，一括確認法，個別確認法よりも短い対話で必要な情報を伝達できることが示された．

今後の課題としては以下のことがある．

- 本章では，確認対話における音声対話システムの対話行為として，確認行為と情報要求行為を取り上げた．しかし，確認対話中のシステムの対話行為としては，ユーザ発話内容の曖昧さの解消や，システムの語彙の範囲などをユーザに教示するといった行為も考えることができる．そういった本章では扱わなかったシステムの対話行為も考慮に入れた上で，対話の長さを最小化するための対話制御法を開発することは今後の課題である．
- 本章では，できるだけ短い対話でユーザが必要とする情報を伝達するという対話の効率性を向上させるための対話制御法に焦点をあてた．しかし，5.1節で述べたように，音声対話システムの有用性は，対話の効率だけではなく，ユーザ満足度も考慮する必要がある．今後の課題としては，本章で議論した対話の長さを最小化するという尺度だけでなく，ユーザ満足度を最大化するという尺度も組み合わせて，対話制御を行う方法を開発していくことが考えられる．

第 6 章

結論

6.1 本研究の要約

本研究では、主導権混在型の音声対話システムに焦点を当て、システムの各構成部を適切に連動させることにより、システムとユーザの間に円滑な音声対話コミュニケーションを実現するという目的の下に、システムの構成法について論じた。

音声対話システムは、発話理解部、発話生成部、対話制御部から構成される。本論文では、システムが自然で効率的な対話を行うために、構成部がどのように連動して働けばよいかという観点から、語用論的条件の充足に基づく発話理解、漸次的発話生成、効率的対話制御という3つの研究テーマに取り組んだ。

第一に、日本語対話において、待遇表現、受給表現、特定の文末形式の使用に際して課せられる語用論的条件を文脈の下に充足させることにより、対話登場人物を指示するゼロ代名詞の指示対象を決定する発話理解法を示した。ATR 自動翻訳電話研究所で収集されたキーボード対話において、提案法は82%のゼロ代名詞を正しく解析できた。提案法は、従来法が前提としているようなドメインの行為やイベントの因果関係に関する膨大な量の知識に頼ることなく、対話登場人物を指示するゼロ代名詞の指示対象を決定できるという利点がある。提案法は、ユーザ発話に含まれるゼロ代名詞の指示対象を文脈情報を使って決定するだけでなく、ユーザ発話によって導入される新たな文脈情報を取り出すことができる。発話理解部と発話生成部が文脈情報を共有して連動することにより、ユーザ発話やシステム発話によって導入される文脈情報を、後続の対話におけるユーザ発話の理解

だけでなく、適切なシステム発話生成に役立てることが可能となり、自然な対話を実現できる。

第二に、音声対話システムが、ユーザからのアイツチや割り込み発話に即座に対処し、臨機応変に話の進め方を変更しながら、自然な発話を生成することを可能とする漸次的発話生成法を示した。音声対話では書き言葉の文よりも小さな単位がコミュニケーションの単位となる。したがって、書き言葉の文を単位として発話を生成する従来法は、システム発話生成途中のユーザからのアイツチや割り込み発話に即座に対処しながら発話を生成することは難しい。提案法は話し言葉特有の小さな発話単位で漸次的に発話を生成する。小さな発話単位で段階的に情報を伝達することにより、システム発話途中のユーザ応答に即座に対処することが可能となる。加えて、発話内容を組み立てながら発話を生成することにより、発話内容が最終的に決定されるのを待たずに発話を開始できるので、発話すべきときに即座に発話を開始することが容易となる。提案法では、人同士の対話を書き起した対話データの分析と言語学的な知見から導いた協調的対話原則にしたがって、発話生成部が発話理解部、対話制御部と適切に連動することにより、ユーザのアイツチ、復唱発話、割り込み発話、沈黙に即座に対処し、ユーザ意図に応じて臨機応変に話の進め方を変更する。このことにより、システムは、対話の中でやり取りされる情報についてユーザとの間で共通理解を維持しながら、自然な発話を生成することが可能となる。

第三に、音声認識精度とシステムのデータベース内容に基づいて、できるだけ短い対話でユーザが必要とする情報を伝達することを可能とする対話制御法として、デュアルコスト法、試行型デュアルコストを示した。デュアルコスト法は、ユーザの承認発話により承認済みとなった問い合わせ内容のみが正しいとして応答を生成する方法であり、試行型デュアルコスト法は、デュアルコスト法を発展させ、承認済みでない問い合わせ内容も正しいと仮定した上で応答を生成する方法である。両提案法は、対話の長さを音声認識精度に依存する確認コストとデータベース内容に依存する情報伝達コストの和として推定する。確認コストと情報伝達コストは相反する関係にある。2つのコストの和を最小にするように対話の各時点におけるシステム行動を選択することにより、効率的な対話が実現できる。対話制御部は、対話の各時点において、システム理解状態を随時更新する発話理解部と、

予測されるシステム応答の長さを計算する発話生成部と連動しながら，コストの和が最小となるような対話を導くシステム行動を選択する．結果として，従来法では回避できなかった不必要な確認を回避し，効率的な対話を実現することが可能となる．シミュレーション対話実験と被験者実験の結果により，提案法が従来法よりも短い対話でユーザが必要とする情報を伝達できることが示された．

6.2 今後の展望

音声対話システムの構成部を適切に連動させることによって円滑な音声対話を実現するという観点から今後の展望を述べる．

語用論的条件の充足に基づく発話理解法は，文脈情報に基づいてユーザ発話内容を理解するとともに，ユーザ発話から新規な文脈情報を導き出し，後続対話におけるユーザ発話理解やシステム発話生成に役立てることができる．今後の展望として，本研究で利用した対話登場人物間の社会関係，話し手の視点，情報のなわばりに関する文脈情報や，従来より利用されてきた対話参加者のプランや注視状態，代名詞と指示対象候補の間の文間距離といった文脈情報に基づいて，ゼロ代名詞や省略表現を含む簡潔なシステム発話を生成するための方法を構築することが考えられる．文脈に依存する簡潔なシステム発話の生成は，効率的な対話の実現に貢献する．本論文では，効率的な対話を実現するための対話制御法としてデュアルコスト法を提案したが，システム応答の表層言語形式まで制御して，対話を効率化するということは検討しなかった．デュアルコスト法と，文脈に依存する簡潔なシステム発話の生成を組み合わせることによって，システム応答の表層言語形式の選択まで含めて対話を制御することができれば，より効率的な対話を実現することが期待できる．そのためには，対話制御部と発話生成部がより密接に連動することが必要である．

漸次的発話生成法では，システム発話途中のユーザからのアイツチ，復唱発話，割り込み発話に対処する際に，ユーザ音声の認識誤りについて考慮しなかった．しかし，システム発話途中に起きるユーザ応答が定型的な言語表現として実現されると仮定したとしても，ユーザ音声の認識誤りが生じる可能性はある．今後の展望としては，システム発話途中に起きるユーザ応答の認識・理解に誤りが起きる場

合にも対処できるように，提案法を発展させていくことが考えられる．そのためには，実施中の発話プランを文脈情報として利用して，ユーザ発話の認識・理解結果の曖昧さを解消する方法や，試行型デュアルコスト法で議論したように，音声の認識誤りのためにユーザ応答に誤った対処を行うリスクも考慮した上で，ユーザ応答への対処法を選択するための方法を構築することが必要である．

デュアルコスト法，試行型デュアルコスト法では，システム情報伝達フェーズにおいて，ユーザからの割り込み発話を考慮していない．しかし，試行型デュアルコスト法のように，ユーザが求めている情報を伝達するという誤ったシステム応答を生成する可能性がある場合には，システム応答の途中でユーザが割り込むことを許すことによって，より効率的な対話を実現できると考えられる．今後の展望としては，試行型デュアルコスト法と漸次的発話生成法を組み合わせることで，システム応答途中でのユーザ割り込みも考慮して対話のコストを計算するように，試行型デュアルコスト法を発展させることが考えられる．

本論文で論じたように，音声対話システムでは，複数の構成部が適切に連動することによって，文脈情報を利用した発話理解，漸次的発話生成，効率的対話制御といった様々な機能が実現される．本論文においては，各機能の効果に対する評価を別々に行ったが，今後は，複数の機能を同時に備えたシステムによって，システム全体のパフォーマンスを評価し，システム全体のパフォーマンスに対して各機能が果たす役割を分析することが考えられる．そうすることによって，円滑な音声対話インタフェースの実現に向けて，より効果的に音声対話システム研究を進めていくことが可能となる．

謝辞

本論文をまとめるにあたり，御指導，御教示を賜った北陸先端科学技術大学院大学情報科学研究科の島津明教授に深く感謝致します．貴重な御意見，御助言を頂きました北陸先端科学技術大学院大学情報科学研究科の東条敏教授，同研究科の白井清昭助教授ならびに鳥澤健太郎助教授，東京大学社会情報研究所の石崎雅人助教授に心より御礼申し上げます．

本論文の第 3 章の研究は，ATR 自動翻訳電話研究所において行われたものです．研究の機会を与えてくださるとともに，数々の御指導を頂きました樽松明氏（現在，電気通信大学），相澤輝昭氏（現在，広島市立大学），飯田仁氏（現在，東京工科大学），森元逞氏（現在，福岡大学）に厚く御礼申し上げます．また，研究を進めるにあたって，多くの御助言を頂いた小暮潔氏（現在，ATR メディア情報科学研究所）に心より御礼申し上げます．

第 4 章と第 5 章の研究は，NTT 基礎研究所と NTT コミュニケーション科学基礎研究所において行われたものです．研究の機会を与えてくださるとともに，数々の御指導を頂きました石井健一郎氏（現在，名古屋大学），萩田紀博氏（現在，ATR メディア情報科学研究所），村瀬洋氏（現在，名古屋大学），牧野昭二氏，川端豪氏（現在，関西学院大学），相川清明氏（現在，東京工科大学）に深く感謝致します．本論文をまとめるにあたって御指導頂きました北陸先端科学技術大学院大学の島津明教授には NTT 在籍中にも数々の御指導を頂きました．重ねて御礼申し上げます．

本研究を進めるにあたっては，NTT コミュニケーション科学基礎研究所の数多くの方々に御討論頂くとともに，多くの御意見と御協力を頂きました．NTT コミュニケーション科学基礎研究所の中野幹生氏には，自然言語処理ならびに音声対話システムの多岐にわたる研究課題について，多くの時間を割いて御討論頂くとともに，貴重な御意見と御協力を数多く頂きました．同研究所の杉山聡氏，安田宜仁氏，高橋康博氏には，発話生成と対話制御の研究課題について，日頃より熱心

に御討論頂くとともに，多くの面で御協力を頂きました．同研究所の川森雅仁氏（現在，NTT サイバーソリューション研究所），平沢純一氏（現在，NTT サイバースペース研究所），宮崎昇氏（現在，NTT サイバースペース研究所），石塚健太郎氏，東中竜一郎氏，須藤克仁氏，Matthias Denecke 氏には様々な観点から御討論頂きました．ここに深く感謝の意を表します．

参考文献

- [1] James F. Allen and C. Raymond Perrault. Analyzing intention in utterances. *Artificial Intelligence*, Vol. 15, pp. 143–178, 1980.
- [2] Liliana Ardissono. *Dynamic User Modeling and Plan Recognition in Dialogue*. PhD thesis, Università di Torino, 1996.
- [3] Herald Aust, Martin Oerder, Frank Seide, and Volker Steinbiss. The Philips automatic train timetable information system. *Speech Communication*, Vol. 17, pp. 249–262, 1994.
- [4] Jon Barwise. *The Situation in Logic – CSLI Lecture Notes Number 17*. Stanford: CSLI Publications, 1989.
- [5] Jon Barwise and John Perry. *Situations and Attitudes*. MIT Press, 1983.
- [6] Daniel G. Bobrow, Ronald M. Kaplan, Martin Kay, Donald A. Norman, Henry Thompson, and Terry Winograd. GUS, a frame driven dialog system. *Artificial Intelligence*, Vol. 8, No. 2, pp. 155–173, 1977.
- [7] Susan E. Brennan, Marilyn Walker Friedman, and Carl J. Pollard. A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics (ACL-87)*, pp. 155–162, 1987.
- [8] Sandra Carberry. Pragmatics-based approach to understanding intersentential ellipsis. In *Proceedings of the 23rd Annual Meeting for Computational Linguistics*, pp. 188–197, 1985.
- [9] Jean Carletta, Richard Caley, and Stephen Isard. A system architecture for simulating time-constrained, language production. Research Paper

HCRC/RP-43, University of Edinburgh, 1993.

- [10] Alison Cawsey. Generating interactive explanations. In *Proceedings of 9th National Conference on Artificial Intelligence (AAAI-91)*, pp. 86–91, 1991.
- [11] Jennifer Chu-Carroll. MIMIC: an adaptive mixed initiative spoken dialogue system for information queries. In *Proceedings of the 6th Applied Natural Language Processing (ANLP-2000)*, pp. 97–104, 2000.
- [12] Herbert H. Clark. *Using language*. The Cambridge University Press, 1996.
- [13] Herbert H. Clark and Edward F. Schaefer. Contributing to discourse. *Cognitive Science*, Vol. 13, pp. 259–294, 1989.
- [14] Ronald Cole, David G. Novick, Mark Fanty, Pieter Vermeulen, Stephen Sutton, Dan Burnett, and Johan Schalkwyk. A prototype voice-response questionnaire for the US census. In *Proceedings of the 3rd International Conference on Spoken Language Processing (ICSLP 94)*, pp. 683–686, 1994.
- [15] Stephen Crain and Mark Steedman. On not being led up the garden path: the use of context by the psychological syntax processor. In David Dowty, Lauri Karttunen, and Arnold Zwicky, editors, *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives, ACL Studies in Natural Language Processing*, pp. 320–358. Cambridge University Press, 1985.
- [16] David Crystal. Neglected grammatical factors in conversational English. In Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik, editors, *Studies in English Linguistics: For Randolph Quirk*, pp. 153–166. Longman, 1979.
- [17] Koenraad De Smedt and Gerard Kempen. Segment grammar: a formalism for incremental sentence generation. In Cécile L. Paris, William R. Swartout, and William C. Mann, editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, pp. 329–349. Kluwer Academic Publishers, 1991.

- [18] 堂坂浩二. 対話参加者の心的状態に関する制約に基づく発話解釈モデル. コンピュータソフトウェア (日本ソフトウェア科学会論文誌), Vol. 6, No. 4, pp. 16–29, 1989.
- [19] Kohji Dohsaka. Identifying the referents of zero-pronouns in Japanese dialogues. In *Proceedings of the 9th European Conference on Artificial Intelligence (ECAI-90)*, pp. 240–245, 1990.
- [20] Kohji Dohsaka. Exploiting reference interaction in resolving temporal reference. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING-94)*, pp. 1138–1144, 1994.
- [21] 堂坂浩二. 語用論的条件の解釈に基づく日本語ゼロ代名詞の指示対象同定. 情報処理学会論文誌, Vol. 35, No. 5, pp. 768–778, 1994.
- [22] 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.
- [23] Kohji Dohsaka and Akira Shimazu. A computational model of incremental utterance production in task-oriented dialogues. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING-96)*, pp. 304–309, 1996.
- [24] Kohji Dohsaka and Akira Shimazu. A system architecture for spoken utterance production in collaborative dialogue. In *Working Notes on IJCAI-97 Workshop on Collaboration, Cooperation and Conflict in Dialogue Systems*, pp. 25–31, 1997.
- [25] 堂坂浩二, 安田宜仁, 相川清明. システム知識制限下での効率的音声対話制御. 自然言語処理, Vol. 9, No. 1, pp. 43–63, 2002.
- [26] Kohji Dohsaka, Norihito Yasuda, and Kiyoaki Aikawa. Efficient spoken dialogue control depending on the speech recognition rate and system’s database. In *Proceedings of the 8th European Conference on Speech Communication and Technology (Eurospeech-2003)*, 2003.

- [27] Kohji Dohsaka, Norihito Yasuda, Noboru Miyazaki, Mikio Nakano, and Kiyoaki Aikawa. An efficient dialogue control method under system's limited knowledge. In *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000)*, Vol. 2, pp. 739–742, 2000.
- [28] 堂坂浩二, 安田宣仁, 宮崎昇, 中野幹生, 相川清明. 音声対話システム「飛遊夢 (ひゅうむ)」. 電子情報通信学会 2001 年総合大会講演論文集, pp. 506–507, 2001.
- [29] Laila Dybkjær, Niels Ole Bernsen, and Hans Dybkjær. A methodology for diagnostic evaluation of spoken human-machine interaction. *International Journal of Human-Computer Studies*, Vol. 48, pp. 605–625, 1998.
- [30] 江原暉将, 金淵培. 確率モデルにおけるゼロ主語の補完. 自然言語処理, Vol. 3, No. 4, pp. 67–86, 1996.
- [31] George Ferguson and James F. Allen. TRIPS: an integrated intelligent problem-solving assistant. In *Proceedings of the 16th National Conference on Artificial Intelligence (AAAI-98)*, pp. 567–572, 1998.
- [32] James Glass and Eugene Weinstein. SpeechBuilder: facilitating spoken dialogue system development. In *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech2001)*, pp. 1335–1338, 2001.
- [33] H. Paul Grice. Logic and conversation. In Peter Cole and Jerry L. Morgan, editors, *Syntax and Semantics Vol.3: Speech Acts*, pp. 43–58. Academic Press, 1975.
- [34] Barbara J. Grosz and Candace L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, Vol. 12, pp. 175–204, 1986.
- [35] M. A. K Halliday. *An Introduction to Functional Grammar*. Edward Arnold, 1994.

- [36] 橋田浩一. 制約と言語. コンピュータソフトウェア, Vol. 4, No. 4, pp. 16–29, 1989.
- [37] Ryuichiro Higashinaka, Mikio Nakano, and Kiyoaki Aikawa. Corpus-based discourse understanding in spoken dialogue systems. In *Proceedings of 41st Annual Meeting of the Association for Computational Linguistics (ACL2003)*, pp. 240–247, 2003.
- [38] Eduard H. Hovy. Automated discourse generation using discourse structure relations. *Artificial Intelligence*, Vol. 63, pp. 341–385, 1993.
- [39] 飯田仁. 自然言語対話の言語運用特性と対話処理の研究課題. 人工知能学会誌, Vol. 3, No. 4, pp. 445–452, 1988.
- [40] 飯田仁. 対話翻訳と高度自然言語処理. 人工知能学会誌, Vol. 6, No. 3, pp. 328–337, 1991.
- [41] Megumi Kameyama. A property-sharing constraint in centering. In *Proceedings of the 24th Annual Meetings for Computational Linguistics*, pp. 200–206, 1988.
- [42] 神尾昭雄. 談話における視点. 日本語学, Vol. 4, No. 12, pp. 10–21, 1985.
- [43] Candace Kamm, Shrikanth Narayanan, Dawn Dutton, and Russell Ritenour. Evaluating spoken dialogue systems for telecommunication services. In *Proceedings of the 5th European Conference on Speech Communication and Technology (Eurospeech-97)*, pp. 2203–2206, 1997.
- [44] 川森雅仁, 島津明. 対話における発話交代の分析. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-95-73, pp. 31–38, 1996.
- [45] Gerard Kempen and Edward Hoenkamp. An incremental procedural grammar for sentence formulation. *Cognitive Science*, Vol. 11, pp. 201–258, 1987.

- [46] Kiyoshi Kogure. Parsing Japanese spoken sentences based on HPSG. In *Proceedings of International Workshop on Parsing Technologies*, pp. 132–141, 1989.
- [47] 駒谷和範, 河原達也. 音声認識結果の信頼度を用いた効率的な確認・誘導を行う対話管理. 情報処理学会論文, Vol. 43, No. 10, pp. 3078–3086, 2002.
- [48] Susumu Kuno. Empathy and syntax. *Linguistic Inquiry*, Vol. 8, No. 4, pp. 627–672, 1977.
- [49] Willem J. M. Levelt. *Speaking: From Intentions to Articulation*. The MIT Press, 1989.
- [50] Esther Levin, Roberto Pieraccini, and Wieland Eckert. Using Markov decision process for learning dialogue strategies. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP98)*, 1998.
- [51] Diane J. Litman. Understanding plan ellipsis. In *Proceedings of the 4th National Conference on Artificial Intelligence*, pp. 619–624, 1986.
- [52] Diane J. Litman and James F. Allen. A plan recognition model for sdialogues in conversation. *Cognitive Science*, Vol. 11, pp. 163–200, 1987.
- [53] Diane J. Litman, Michael S. Kearns, Satinder Singh, and Marilyn A. Walker. Automatic optimization of dialogue management. In *Proceedings the 18th International Conference on Computational Linguistics (COLING-2000)*, 2000.
- [54] Donald W. Loveland. A unifying view of some linear Herbrand procedures. *Journal of the Association for Computing Machinery (JACM)*, Vol. 19, No. 2, pp. 366–384, 1972.
- [55] William C. Mann and Sandra A. Thompson. Rhetorical structure theory: towards a functional theory of text. *Text*, Vol. 8, No. 3, pp. 243–281, 1988.
- [56] 益岡隆志, 田窪行則. 基礎日本語文法 - 改定版 -. くろしお出版, 1992.

- [57] Katheleen F. McCoy. Generating context-sensitive responses to object-oriented misconceptions. *Artificial Intelligence*, Vol. 41, pp. 157–195, 1989.
- [58] Kathleen R. McKeown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press, 1985.
- [59] 水谷静夫. 待遇表現の仕組み. 林四郎, 荻野綱男, 田中幸子, 樺島忠夫 (編), 運用 I, pp. 158–178. 朝倉書店, 1983.
- [60] 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.
- [61] 中川聖一. 確率モデルによる音声認識. (社) 電子情報通信学会, 1988.
- [62] 中岩浩巳, 池原悟. 語用論的・意味論的制約を用いた文内照応解析. 自然言語処理, Vol. 3, No. 4, pp. 49–65, 1996.
- [63] Mikio Nakano. Constraint projection: an efficient treatment of disjunctive feature descriptions. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics (ACL-91)*, pp. 307–314, 1991.
- [64] Mikio Nakano, Kohji Dohsaka, Noboru Miyazaki, Junichi Hirasawa, Masafumi Tamoto, Masahito Kawamori, Akira Sugiyama, and Takeshi Kawabata. Handling rich turn-taking in spoken dialogue systems. In *Proceedings of the 6th European Conference on Speech Communication and Technology (Eurospeech99)*, pp. 1167–1170, 1999.
- [65] Mikio Nakano, Noboru Miyazaki, Junichi Hirasawa, Kohji Dohsaka, and Takeshi Kawabata. Understanding unsegmented user utterances in real-time spoken dialogue systems. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*, pp. 200–207, 1999.

- [66] Yasuhisa Niimi and Yutaka Kobayashi. Dialog control strategy based on the reliability of speech recognition. In *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP96)*, 1996.
- [67] Cécile L. Paris. *User Modelling in Text Generation*. Francis Pinte, 1993.
- [68] Thomas Pechmann. Incremental speech production and referential overspecification. *Linguistics*, Vol. 27, pp. 89–110, 1989.
- [69] Jeremy Peckham. A new generation of spoken language systems: results and lessons from the SUNDIAL project. *Proceedings of the 3rd European Conf. Speech Communication and Technology (Eurospeech-93)*, pp. 33–40, 1993.
- [70] Massimo Poesio. A model of conversation processing based on micro conversational events. In *Proceedings of the Annual Meeting of the Cognitive Science*, 1995.
- [71] Bhavani Raskutti and Ingrid Zukerman. Query and response generation during information-seeking interactions. In *Proceedings of the 4th International User Modelling Conference*, pp. 25–30, 1994.
- [72] Ehud Reiter and Chris Mellish. Optimizing the costs and benefits of natural language generation. In *Proceedings of 13th International Joint Conference on Artificial Intelligence (IJCAI-89)*, pp. 1164–1169, 1993.
- [73] Norbert Reithinger. POPEL – a parallel and incremental natural language generation. In Cécile L. Paris, William R. Swartout, and William C. Mann, editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, pp. 179–199. Kluwer Academic Publishers, 1991.
- [74] Norbert Reithinger. The performance of an incremental generation for multimodal dialog contributions. In Robert Dale, Eduard Hovy, Dietmar Rösner, and Oliviero Stock, editors, *Aspects of Automated Natural Language Generation*, pp. 263–276. Springer-Verlag, 1992.

- [75] Nicholas Roy, Joelle Pineau, and Sebastian Thrun. Spoken dialogue management using probabilistic reasoning. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL2000)*, 2000.
- [76] Stuart Russel and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 1995.
- [77] John R. Searle. *Expression and Meaning*. Cambridge University Press, 1979.
- [78] 島津明, 川森雅仁. 対話データの表記法. 情報処理学会第44回全国大会講演論文集 (3) 1N-5, pp. 99–100, 1992.
- [79] Wee Meng Soon, Hwee Tou Ng, and Daniel Chung Yon Lim. A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, Vol. 27, No. 4, pp. 521–544, 2001.
- [80] Anna-Brita Stenström. *An Introduction to Spoken Interaction*. Longman, 1994.
- [81] Michael Strube and Christoph Muller. A machine learning approach to pronoun resolution in spoken dialogue. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL-2003)*, pp. 168–175, 2003.
- [82] Alan M. Turing. Computing machinery and intelligence. *Mind*, Vol. 59, No. 236, pp. 433–460, 1950.
- [83] Peter van Beek and Robin Cohen. Resolving plan ambiguity for cooperative response generation. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91)*, pp. 938–944, 1991.
- [84] Marilyn A. Walker, Diane J. Litman, Candace A. Kamm, and Alicia Abella. Evaluating spoken dialogue agents with PARADISE: two case studies. *Computer Speech and Language*, Vol. 12, No. 3, pp. 317–347, 1998.

- [85] Marilyn A. Walker and Owen Rambow. The role of cognitive modeling in achieving communicative intentions. In *Proceedings of the 7th International Conference on Natural Language Generation*, 1994.
- [86] Marilyn Walker, Masayo Iida, and Sharon Cote. Japanese discourse and the process of centering. *Computational Linguistics*, Vol. 20, No. 2, pp. 193–232, 1994.
- [87] Marilyn Walker and Steve Whittaker. Mixed initiative in dialogue: an investigation into discourse segmentation. In *Proceedings of 28th Annual Meeting of the Association for Computational Linguistics*, pp. 70–78, 1990.
- [88] 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.
- [89] 安田宜仁, 堂坂浩二, 相川清明. タスク適応型高効率対話制御法. 情報処理学会研究会報告, 音声言語情報処理研究会, SLP-35-13, pp. 73–78, 2001.
- [90] Kei Yoshimoto. Identifying zero pronouns in Japanese dialogue. In *Proceedings of the 13th International Conference on Computational Linguistics (COLING-88)*, pp. 779–784, 1988.
- [91] Victor Zue, Stephanie Seneff, Joseph Polifroni, Michael Phillips, Christine Pao, David Goodine, David Goddeau, and James Glass. PEGASUS: a spoken dialogue interface for on-line air travel planning. *Speech Communication*, Vol. 15, pp. 331–340, 1994.

研究業績

査読付きの論文

1. 堂坂浩二, 安田宜仁, 相川清明. システム知識制限下での効率的音声対話制御. 自然言語処理, Vol. 9, No. 1, pp. 43–63, 2002.
2. 杉山聡, 堂坂浩二, 川端豪. 音声対話によるテキスト内容の伝達方法. 情報処理学会論文誌, Vol. 41, No. 6, pp. 1883–1894, 2000.
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. 語用論的条件の解釈に基づく日本語ゼロ代名詞の指示対象同定. 情報処理学会論文誌, Vol. 35, No. 5, pp. 768–778, 1994.
5. 堂坂浩二. 対話参加者の心的状態に関する制約に基づく発話解釈モデル. コンピュータソフトウェア (日本ソフトウェア科学会論文誌), Vol. 6, No. 4, pp. 16–29, 1989.

査読付きの国際会議

1. Kohji Dohsaka, Norihito Yasuda, and Kiyoaki Aikawa. Efficient spoken dialogue control depending on the speech recognition rate and system's database. In *Proceedings of the 8th European Conference on Speech Communication and Technology (Eurospeech-2003)*, 2003.
2. Yasuhiro Takahashi and Kohji Dohsaka. An efficient dialogue control method using decision-tree-based estimation of out-of-vocabulary word attributes.

- In *Proceedings of International Conference on Spoken Language Processing (ICSLP2002)*, pp. 813–816, 2002.
3. Norihito Yasuda, Kohji Dohsaka, and Kiyoaki Aikawa. Spoken dialogue control based on a turn-minimization criterion depending on the speech recognition accuracy. In *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialog (SIGdial2001)*, pp. 210–213, 2001.
 4. Kohji Dohsaka, Norihito Yasuda, Noboru Miyazaki, Mikio Nakano, and Kiyoaki Aikawa. An efficient dialogue control method under system’s limited knowledge. In *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000)*, Vol. 2, pp. 739–742, 2000.
 5. Mikio Nakano, Kohji Dohsaka, Noboru Miyazaki, Junichi Hirasawa, Masafumi Tamoto, Masahito Kawamori, Akira Sugiyama, and Takeshi Kawabata. Handling rich turn-taking in spoken dialogue systems. In *Proceedings of the 6th European Conference on Speech Communication and Technology (Eurospeech99)*, pp. 1167–1170, 1999.
 6. Kohji Dohsaka and Akira Shimazu. A system architecture for spoken utterance production in collaborative dialogue. In *Working Notes on IJCAI-97 Workshop on Collaboration, Cooperation and Conflict in Dialogue Systems*, pp. 25–31, 1997.
 7. Kohji Dohsaka and Akira Shimazu. A computational model of incremental utterance production in task-oriented dialogues. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING-96)*, pp. 304–309, 1996.
 8. Kohji Dohsaka. Exploiting reference interaction in resolving temporal reference. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING-94)*, pp. 1138–1144, 1994.

9. Kohji Dohsaka. Identifying the referents of zero-pronouns in Japanese dialogues. In *Proceedings of the 9th European Conference on Artificial Intelligence (ECAI-90)*, pp. 240–245, 1990.

査読なしの発表

1. 堂坂浩二, 安田宜仁, 相川清明. 試行型対話戦略による音声対話システムの確認発話削減. 言語処理学会第9回年次大会発表論文集, pp. 63–66, 2003.
2. 安田宜仁, 堂坂浩二, 相川清明, 上野晋一. シングルドメインシステムの統合による複数ドメイン音声対話システム. 情報処理学会研究報告, 音声言語情報処理研究会, SLP-45-20, pp. 119–124, 2003.
3. 堂坂浩二, 安田宜仁, 相川清明. 情報伝達効率最大化に基づく音声対話制御法. 言語処理学会第8回年次大会発表論文集, pp. 260–263, 2002.
4. 安田宜仁, 堂坂浩二, 相川清明. 2つの認識文法を用いた主導権混合型対話制御. 情報処理学会研究会報告, 音声言語情報処理研究会, SLP-40-21, pp. 127–132, 2002.
5. 安田宜仁, 堂坂浩二, 相川清明. 2つの認識文法を用いた主導権混合型対話制御. 日本音響学会講演論文集, Vol.I, pp. 77–78, 2002.
6. 高橋康博, 堂坂浩二, 相川清明. 決定木を用いた未知語属性推定による効率的対話方法. 電子情報通信学会技術研究報告, 音声研究会, SP2002-34, pp. 61–66, 2002.
7. 高橋康博, 堂坂浩二, 相川清明. 決定木を用いた未知語属性推定による効率的対話方法. 言語処理学会第8回年次大会発表論文集, pp. 256–259, 2002.
8. 堂坂浩二, 安田宜仁, 相川清明. 対話コスト最小化に基づく効率的対話制御. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC2001-34, pp. 93–100, 2001.

9. 堂坂浩二, 相川清明. 対話コスト最小化原理に基づく対話制御. 言語処理学会第7回年次大会発表論文集, pp. 518-521, 2001.
10. 堂坂浩二, 安田宣仁, 宮崎昇, 中野幹生, 相川清明. 音声対話システム「飛遊夢(ひゅうむ)」. 電子情報通信学会2001年総合大会講演論文集, pp. 506-507, 2001.
11. 安田宣仁, 堂坂浩二, 相川清明. タスク適応型高効率対話制御法. 情報処理学会研究会報告, 音声言語情報処理研究会, SLP-35-13, pp. 73-78, 2001.
12. 安田宣仁, 堂坂浩二, 相川清明. タスク適応型高効率対話制御法. 言語処理学会第7回年次大会発表論文集, pp. 526-529, 2001.
13. 安田宣仁, 堂坂浩二, 相川清明. 確認対象での認識性能を考慮する効率的対話制御法. 日本音響学会講演論文集, Vol.I, pp. 51-52, 2001.
14. 高橋康博, 堂坂浩二, 相川清明. 音声対話における実例に基づく未知語属性推定. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC2001-22~35, pp. 101-106, 2001.
15. 高橋康博, 堂坂浩二, 相川清明. 音声対話における実例に基づく未知語属性推定. 言語処理学会第7回年次大会発表論文集, pp. 522-525, 2001.
16. 堂坂浩二, 中野幹生, 宮崎昇, 安田宣仁, 相川清明. システム制限知識下における効率的対話制御. 情報処理学会研究会報告, 音声言語情報処理研究会, 2000-SLP-33, pp. 49-54, 2000.
17. 堂坂浩二, 中野幹生, 宮崎昇, 安田宣仁, 相川清明. 制限知識下における効率的対話制御. 言語処理学会第6回年次大会発表論文集, pp. 340-343, 2000.
18. 杉山聡, 堂坂浩二, 相川清明. 対話履歴を利用した協調的応答生成. 言語処理学会第6回年次大会発表論文集, pp. 227-230, 2000.
19. 堂坂浩二. 音声対話システムと人間の対話. CmCC研究会, 第1回シンポジウム, 1999.

20. 堂坂浩二, 川端豪. 話しことば対話における協同理解戦略. 言語処理学会第5回年次大会発表論文集, pp. 263-266, 1999.
21. 中野幹生, 堂坂浩二, 宮崎昇, 平澤純一, 田本真詞, 川森雅仁, 杉山聡, 川端豪. 柔軟な話者交代を行う音声対話システム dug-1. 言語処理学会第5回年次大会発表論文集, pp. 161-164, 1999.
22. 堂坂浩二, 川端豪, 島津明. 複数の対話ドメインにおける協調的対話原則の分析. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-97-58, pp. 25-32, 1998.
23. 杉山聡, 堂坂浩二, 川端豪. 話しことば対話によるテキスト内容の伝達. 情報処理学会研究会報告, 音声言語情報処理研究会, 98-SLP-21, pp. 83-93, 1998.
24. 中野幹生, 堂坂浩二, 宮崎昇, 平澤純一, 田本真詞, 川森雅仁, 杉山聡, 川端豪. Tv 番組の録画予約を受け付ける実時間音声対話システム. 音声言語情報処理研究会, 情報処理研究報告, 98-SLP-22, pp. 41-42, 1998.
25. 堂坂浩二, 島津明. 話し言葉対話コーパスにおける協調的対話原則の分析. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-97-1, pp. 1-8, 1998.
26. 堂坂浩二, 島津明. 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. 言語処理学会第3回年次大会発表論文集, pp. 501-504, 1997.
27. 川森雅仁, 島津明, 堂坂浩二, 中野幹生. 対話処理のためのコーパス作成. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-97-5, pp. 31-36, 1997.
28. 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, 言語理解とコミュニケーション研究会, NLC-96-32, pp. 9-16, 1996.
29. 堂坂浩二, 島津明. 談話構造を利用した漸次的発話生成. 言語処理学会第2回年次大会発表論文集, pp. 225-228, 1996.

30. 島津明, 小暮潔, 川森雅仁, 堂坂浩二, 中野幹生. 対話処理システムにおける内的コミュニケーション. 言語処理学会第2回年次大会発表論文集, pp. 333–336, 1996.
31. 堂坂浩二, 島津明. 実時間対話における漸次的発話のプランニングモデル. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-95-40, pp. 1–8, 1995.
32. 堂坂浩二, 島津明. 時間制限下における漸次的発話生成. 情報処理学会第51回(平成7年後期)全国大会講演論文集, pp. 107–108, 1995.
33. 島津明, 小暮潔, 堂坂浩二, 中野幹生. 対話処理のための分散協調モデル. 言語処理学会第1回年次大会発表論文集, pp. 81–84, 1995.
34. 堂坂浩二, 島津明. 時間制限下における漸次的発話生成. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-94-40, pp. 9–16, 1994.
35. 堂坂浩二. 参照の相互作用に基づく時の参照の解釈. 電子情報通信学会技術研究報告, 言語理解とコミュニケーション研究会, NLC-93-64, pp. 7–14, 1994.
36. 堂坂浩二. ゼロ代名詞を取り巻く環境. 情報処理学会第44回(平成4年前期)全国大会講演論文集, pp. 195–196, 1992.
37. 堂坂浩二. 効率的表現を介した解釈と生成. 情報処理学会研究報告, 自然言語処理研究会, 91-NL-82, 1991.
38. 堂坂浩二. 対話登場人物を指示する日本語ゼロ代名詞の同定. 「談話理解モデルとその応用」シンポジウム, 情報処理学会, pp. 41–52, 1989.
39. 堂坂浩二, 小暮潔. 対話参加者に関するゼロ代名詞の同定. 情報処理学会第39回(平成元年後期)全国大会講演論文集, 1989.

40. 堂坂浩二, 小暮潔, 飯田仁. 対話参加者の心的状態に関する制約を用いた文脈処理手法. 情報処理学会第 38 回 (平成元年前期) 全国大会講演論文集, pp. 251–252, 1989.
41. 堂坂浩二. 対話参加者の心的状態に関する制約を用いた文脈処理手法. ディスコースと形式意味論ワークショップ, ソフトウェア科学会「論理と自然言語」研究会 と (財) 新世代コンピュータ技術開発機構 共催, pp. 123–132, 1989.
42. 堂坂浩二, 飯田仁. 対話における意図構造の表現の枠組. 情報処理学会第 37 回 (昭和 63 年後期) 全国大会講演論文集, pp. 1113–1114, 1988.

解説等

1. 中野幹生, 堂坂浩二. 音声対話システムの言語・対話処理. 人工知能学会誌, Vol. 17, No. 3, pp. 271–278, 2002.
2. 堂坂浩二. 発話環境解析. 国際電気通信基礎技術研究所 (編), ATR 先端テクノロジーシリーズ「自動翻訳電話」, 3.7 章. オーム社, 1992.