

表層表現パターンを用いた対話構造の認識

巖 寺 俊 哲[†] 石 崎 雅 人^{††} 森 元 暉^{†††}

対話中の発話では、先行する発話で述べられたことが前提とされるため省略や代用表現が多用される。このような発話に対して翻訳等の処理を適切に施すためには、先行する発話との関連を認識することが必要である。本論文では、発話間の対応関係を認識することを目的として、対話構造のモデルを提示し、このモデルに基づいて対話構造を認識する手法を提案する。本手法では、文末表現等から構成される表層表現パターンを用いて対話構造を認識する。このパターンは、領域知識に依存せず、様々な領域に適用することが可能である。また、本手法を用いることで、発話ごとに漸次的に発話間の対応関係を認識することが可能となる。本手法による発話間の対応付け結果を6人の被験者による対応付け結果と定量的に比較することで評価した。その結果、オープンデータに関して、被験者6人全員が認定した対応関係に対する本手法による対応関係の再現率は、平均93.1%、被験者の少なくとも1人以上が認定した対応関係に対する本手法による対応関係の適合率は、平均87.4%であった。これは、人間に近い精度で発話の対応関係を認識でき、特に、信頼性が高い対応関係を高い精度で認識し、非対応関係に対する誤認識は少ないことを示しており、本手法が、発話間の対応関係の認識の目的に有効であることが確認できた。

Recognizing an Interactional Structure Using Surface Form Patterns

TOSHIAKI IWADERA,[†] MASATO ISHIZAKI^{††}
and TSUYOSHI MORIMOTO^{†††}

In this paper, we present a model of interactional structures and propose a novel method for incrementally recognizing them based on the model. The method aims at recognizing corresponding relations between utterances such as questions and answers in a dialogue. These relations are important for machine translation and computer dialogue systems to process elliptical and substitutional expressions, because the corresponding relations are one of the most important clues to recover the content of those expressions. An interactional structure is recognized using surface expression patterns characterized by sentence final forms. Since the patterns can be built without using the domain knowledge of dialogues, this method can be applied extensively to various domains. We evaluate our proposed method quantitatively by comparing its recognition results with the judgements by 6 human subjects: The recall rate of the recognition results to the judgements agreed by all human subjects is 93.1% on average; the precision rate of the recognition results to the judgements by at least one of the human subjects is 87.4% on average. The results show that our method can recognize correct corresponding relations and set aside wrong relations between utterances with the accuracy comparable to that of human judgement. This demonstrates that our method can effectively recognize the corresponding relations between utterances in dialogues.

1. はじめに

対話中の発話は、先行する発話で述べられたことが

前提とされるため、省略や代用表現が多用される。省略補完や代用表現の解釈は、関連する発話からの情報を援用することによって可能になる。また、対話中の発話を適切に翻訳するには、各発話を個別に翻訳するのではなく、呼応する先行発話に応じて訳し分ける必要がある。すなわち、対話中の各発話がどのように相互に関係しているかを示す対話構造を認識することが必要である。

対話構造のとらえ方は、処理システムの対話へのかかり方によって異なる。ユーザー-システム間の対話を扱う対話システムの場合は、対話参加者の一方がつねにシステムであり、対話の流れを制御できる。この

[†] NTT 情報通信研究所

NTT Information and Communication Systems Laboratories

^{††} 北陸先端科学技術大学院大学知識科学研究科

School of Knowledge Science, Japan Advanced Institute of Science and Technology

^{†††} 福岡大学工学部

Faculty of Engineering, Fukuoka University

★ 本研究は、著者全員がATR 音声翻訳通信研究所に在籍に行ったものである。

ため、対話構造をシステム自ら構築しつつ処理を進めることができる。これに対して、対話参加者が相互に能動的に発話する人間同士の対話を仲介し、その発話を翻訳する翻訳システム等の場合は、対話の展開を制御することができない。このため、観測された話者の発話から対話構造を認識することが要求される。

本論文では、発話表現間の対応関係を認識することを目的として、対話構造のモデルを提示するとともに、そのモデルに基づいて対話構造を認識する手法を提案する。本手法は、対話の展開を能動的に制御するのではなく、2人の話者間で交わされる発話を中立的にモニタし、発話がなされるごとに漸次的に対話構造を認識する。このため、翻訳システム等にも適用できる。認識には、文末表現等から構成される表層表現パターンと対話特有の現象である話者交替に関する情報を用いる。これらは、領域に依存しない情報であり、様々な領域に適用することができる。さらに、発話表現間の対応関係を得ることにより、省略補完や代用表現の解釈に有用な情報を提供することが可能となる。

以下、2章で、関連研究について述べる。3章で、発話間の対応関係の認識の必要性について述べる。4章で、この発話間の対応関係を得ることを目的として、対話構造のモデルを提示し、このモデルに基づいて、対話構造を認識する手法について述べる。5章で、「旅行会話」タスクにかかわる対話を対象として行った評価実験とその結果について述べる。6章で、結論を述べる。

2. 関連研究

対話構造とその構造をとらえる手法についていくつかの研究がなされている。

Groszら²⁾は、談話中の発話の処理を記述するための枠組みとなる理論を提示している。この理論では、談話の構造は、相互に関係する3種の要素、言語構造 (linguistic structure)、意図構造 (intentional structure)、注視状況 (attentional structure) から構成される。このような構造が、手がかり語、参照表現、割込みのような談話にかかわる現象を説明するのに重要であることを示している。この理論に基づく談話セグメントは、一貫したゴールを持つ部分を1つの単位としている。しかし、我々が求めようとしている対話構造の単位は、質問-応答等の発話表現間の対応関係に基づくものであり、より小さな単位となる。

Groszら²⁾の談話構造に基づいて、談話セグメントに関する研究が行われている。Hirschbergら⁴⁾は、読み上げられたAPニュース記事からなるコーパスを用

いて、談話セグメントと韻律情報が相互に関係していることを定量的に示している。Groszら³⁾は、APニュース記事を対象に談話セグメント境界の一致度を被験者間で比較している。

Litmanら¹³⁾、Passonneauら⁷⁾は、3つの言語的手がかり (参照名詞句、手がかり語、ポーズ) を用いて談話セグメント境界を予測する手法を提示し、コーパスに対して被験者が割り当てたセグメントと定量的に比較、評価している。

上記の研究では、対象としたデータは、独話 (モノログ) である。独話では、話者は1人であることから、独話中の表現はすべて同一の話者により発話されたものである。これに対して、対話は、2人の話者がかかわっており、2人の話者が相互に相手の発話に応じて発話した表現から構成されている。本研究では、対話を対象としており、文末表現などの発話の表層表現から得られる情報とともに、2人の話者がかかわっている対話特有の現象である話者交替に関する情報を利用している。

Allen¹⁾は、プランの枠組みで発話を解析し協調的な応答をするモデルを示した。このモデルによれば、質問されたことよりも多くの情報を与える応答や、断片的な発話、間接的な発話に協調的に答える応答が可能になる。

飯田ら⁵⁾は、4階層のプラン認識モデルを使った対話理解の手法を提案している。この手法では、インタラクションプラン、コミュニケーションプラン、ドメインプラン、ダイアログプランの4階層からなるプランを利用することで、対話の進展にともなう各発話と話題領域の知識との関係付けを漸次進めていき、対話全体にわたる構造をつくる。この対話構造中で断片的発話の解釈を行うことにより、代用表現、格要素の省略を復元できることを示した。「会議の申し込み」に関する4対話に対して実験を行ったと報告しているが定量的な結果には言及していない。Allen、飯田らの研究で使用しているプランは、領域に依存する知識であり、異なる領域に適用する場合、その領域に応じた知識をあらかじめ人手により作成する必要がある。

本研究では、あらかじめ用意する知識は、特定の領域に依存しない知識である文末表現や手がかり語等の表層表現パターンである。本稿では、「旅行」に関する対話に適用するが、異なる領域への適用も可能であると考えられる。

Litmanら¹²⁾は、実際の対話では、発話間の対応関係が入れ子構造になる場合があり、これに対処することが必要であることを示している。熊本ら¹⁰⁾は、実際

に収集した対話データを基に、入れ子構造の出現割合を調査し、その構造のパターンをモデル化している。さらに、この結果に基づいてユーザの機器利用を支援する対話型支援システムに適した対話処理方式を提案し、対話を階層的に扱うことで、入れ子構造の問題に対処している。この方式では、対話管理と応答生成が一体となっており、対話の一方の話者はつねにシステムであることを想定している。このため、ユーザを問題解決へと誘導するシステム発話を生成することができる。また、ユーザが発話したときのみ、何らかの応答を行う受動的な支援を基本としている。

対話参加者の一方がつねにシステムである対話システムに対して、人間同士の対話を仲介し、その発話を翻訳する音声翻訳システム（たとえば、文献9）、18）、20））のように、対話参加者が両者とも人間である場合がある。このようなシステムでは、対話を能動的に制御できず、外部から対話を観測し処理する方式が要求される。このような研究として、Nagataら¹⁴⁾、Reithingerら¹⁵⁾、Katohら⁸⁾がある。本稿で提案する手法も、対話の展開を能動的に制御するのではなく、2人の話者間で交わされる発話を中立的に外部からモニタすることで対話構造を認識する。

Nagataら¹⁴⁾は、発話を発話行為タイプ IFT (Illocutionary Force Type) で分類し、タイプのマルコフ遷移モデルで対話を表現する。さらに、このモデルを用いて、次発話の発話タイプを予測する。これは、首尾一貫性のない文脈を形成する音声認識結果候補を排除するために使用される。Reithingerら¹⁵⁾は、発話を dialogue act で分類し、先行する対話の dialogue act の系列から次発話の dialogue act を統計的に予測する。この予測結果は、他のモジュールから出力されるプラン認識結果の修正や音声認識時の探索空間の削減に使用される。Katohら⁸⁾は、発話間の結束度を定義し、談話構造を解析し、音声認識処理へ発話間の言語的制約を提供しようとしている。Katohらの手法では、発話を発話意図、動詞、名詞、手がかり語、話題で表現する。発話間の結束性の有無は、手がかり語、発話対上の結束度、話題上の結束度の3つの観点から決定される。結束度は、対話コーパスから統計的に得られた情報を使って計算される。上記のいずれの研究も音声認識等への適用を想定しており、次発話の曖昧性を制約することを主眼としている。これに対して、本研究の主眼は、新たに入力された発話中の省略補完や代用表現の解釈に有用な情報を得るために、関連がある先行発話を特定することである。

対話構造を認識するためには、まず、対話を構成す

る各発話表現を特徴付け、分類することが必要となる。発話表現を分類する研究として熊本ら¹¹⁾、久米ら¹⁹⁾がある。熊本ら¹¹⁾は、「ヘルプシステム」のために12種類の発話意図タイプを設定し、これらをモダリティ情報とパターン-発話意図タイプ変換テーブルとのパターンマッチングによって決定している。この発話意図タイプには、ヘルプシステムの応答アルゴリズムを選択する役割がある。久米ら¹⁹⁾は、「会議問い合わせ」タスクのために7個の発話行為タイプを設定し、文末表現および副詞類などの大局的な制約を用いて発話行為タイプを決定している。原言語から決定された発話行為タイプを目標言語の社会習慣上の伝達様式に関する知識を利用して表現することで、発話意図翻訳に利用する。本研究では、表層表現パターンとそのパターンに照合可能である発話を持ちうる発話行為を網羅的に対応させ記述した対応テーブルを用いて分類する。また、分類時に生じる発話行為の多義は、対話構造を認識する過程で順次解消していくアプローチをとる。

3. 発話間の対応関係

対話中の発話を適切に処理するためには、対応する先行発話を認識し、先行発話の内容を考慮する必要がある。特に、対話中の発話を適切に翻訳するためには、各発話を個別に翻訳するのではなく、呼応する先行発話に応じて訳し分ける必要がある。

たとえば、日本語の典型的な肯定返答表現である「はいそうです」は、“Yes, I am.”、“Yes, it is.”などのように様々な表現に翻訳可能である。これを適切に訳し分けるためには、適切な先行発話を認識し、それに応じて適切な翻訳結果を選択することが必要である。たとえば、表1に示すように、この発話の対応する発話が、“Are you Mr. Suzuki?”の場合は、“Yes, I am.”と翻訳する必要がある。また、否定の発話“‘There isn't a discount this time, is it?’”に対する返答の場合は、この発話に呼応して“‘No, it isn't.’”と翻訳する必要がある。

また、日本語の対話に頻繁に現れる断片的発話や代用的な表現には、先行する対応発話に応じて解釈を替えて翻訳する必要があるものがある。たとえば、表2に示すように、発話「参加料は銀行振込です」は、先行発話に応じて適切に翻訳することが必要である。

上記の例では、対応する先行発話の必要性を示したが、実際の対話中では、対応する発話が連続して発話されるとは限らない。たとえば、表3に示す対話例では、表4に示すように発話番号(1, 2)、(2, 7, 8)、(3, 6)、(4, 5)の発話表現が各々対応している。この

表1 「はいそうです」の翻訳例
Table 1 Translation examples of "Hai, soudesu".

先行発話		「はいそうです」の翻訳結果
Are you Mr. Suzuki?	→	Yes, I am.
That would be a little faster, wouldn't it?	→	Yes, it would.
Is that ferry the same one?	→	Yes, it is.
There isn't a discount this time, is it?	→	No, it isn't.

表2 「参加料は銀行振込です」の翻訳例
Table 2 Translation examples of "Sanka-ryouwa ginkou furikomi desu".

先行発話		翻訳結果
How should I pay the attendance fee by bank-transfer?	→	You should pay the attendance fee by bank-transfer.
How did you pay the attendance fee and hotel expences?	→	I paid the attendance fee by bank-transfer.

表3 対応発話が隣接して現れない対話例
Table 3 A dialogue example including non-adjacent corresponding utterances.

発話番号	発話者	発話表現
1	お客	ばく1枚で禁煙席をお願いします
2	駅員	分かりましたお1名様禁煙席ですねファーストとコーチどちらがよろしいでしょうか
3	お客	お値段はそれぞれいくらくらい違うんですか
4	駅員	片道ですかそれとも往復をお買いになりますか
5	お客	片道で結構です
6	駅員	コーチは片道ですと75ドルになりますファーストですと片道115ドルになります
7	お客	じゃあコーチの方をお願いします
8	駅員	分かりました

表4 対応発話
Table 4 Correspondences between utterances.

発話番号	発話者	発話表現
1	お客	ばく1枚で禁煙席をお願いします
2	駅員	分かりましたお1名様禁煙席ですね
2	駅員	ファーストとコーチどちらがよろしいでしょうか
7	お客	じゃあコーチの方をお願いします
8	駅員	分かりました
3	お客	お値段はそれぞれいくらくらい違うんですか
6	駅員	コーチは片道ですと75ドルになりますファーストですと片道115ドルになります
4	駅員	片道ですかそれとも往復をお買いになりますか
5	お客	片道で結構です

ように、対応発話間に挟み込まれるように、問い返しや確認の発話のやりとりが生起することがあり、発話表現間の対応関係を認識する仕組みは、隣接してい

ない発話表現間の対応関係も扱える必要がある。

さらに、対話システムや翻訳対話システムにおいては、ユーザの発話ごとに処理結果を返す必要があるため、対話がすべて終わってから発話間の対応をとるのではなく、1発話ごとに対応をとる必要がある。

4. 対話構造のモデルと認識

本章では、発話間の対応をとるための対話構造のモデルと対話構造の認識手法について述べる。

4.1 談話分析理論

対話データを観察すると、「ある話者が話しかけ、別の話者がこれに回答する」、あるいは、「話しかけ、回答し、話しかけた話者がさらに回答に対して補足したり評価したりする」、というようなパターンの繰返しを見ることが出来る。また、さらにこれらのパターンが幾重にも入れ子状になって繰り返されている。

談話分析理論では、質問-回答のようなやりとりの対応関係を複数の階層から構成される対話のやりとりの構造としてモデル化している⁶⁾。まず、対話中の発話は、やりとりを考慮して発話行為を拡張した対話行為で特徴付けられる。次に、複数の対話行為が働きかけや応答などの対話の局所構造の中での役割を表現する交換行為にまとめられる。この交換行為のまとまりが交換単位となり、異なる話者の発話の間の関係を表現する。さらに、同一の話題を構成する複数の交換単位は、話題単位を構成する。このように、やりとりから話題のまとまりまで対話をモデル化している。たとえば、文献17)では、ACT, MOVE, TURN, EXCHANGE, TRANSACTIONの5階層からなるモデルが使用されている。ACT, MOVE, EXCHANGE, TRANSACTIONは、上記の対話行為、交換行為、交換単位、話題単位にそれぞれ対応する。

上記の談話分析における階層モデルは、対話を分析するためのモデルであり、対話の展開とともにどのように対話構造が形成されていくかを説明していない。本論文では、上記の階層モデルを参考にして、まず、対話構造のモデルを提示し、次に、発話間の対応関係を認識することを目的とした、対話の展開とともにその構造を漸次的に形成、認識する手法を提案する。

4.2 3階層の対話構造モデル

まず、対話の展開とともにその構造を漸次的に形成、認識するために使用する対話構造のモデルを示す。ここで述べる対話構造は、ACT, MOVE, EXCHANGEの3階層から構成されている。これらは、Stenstrom¹⁷⁾の用語、定義にほぼ対応している。

4.2.1 ACT層

ACT層では、発話行為を表す最小単位(ACT単位)を表現する。対話は、順番に発話権(ターン)が移動し、それぞれターンを持つときに発話がなされる。発話は、単一の発話行為から構成されることもあり、複数のこともある。この発話行為の1つ1つをACT層で表現する。具体的には、Seligman¹⁶⁾が提示している文末表現などの表現と情報伝達、情報要求などの発話行為との関係に基づき、発話を複数の発話行為(ACTラベルと呼ぶ)に分類する。

本研究の目的は、発話表現間の対応関係を認識することであり、Seligman¹⁶⁾が提案しているような詳細な分類は必要ない。しかしながら、音声対話システムや音声翻訳対話システムなどの応用を考えた場合、発話行為の情報は、発話の解釈や翻訳に重要な役割を果たす。ここでは、このような応用システムを念頭において、詳細な発話行為を複数付与している。この複数の発話行為から正しいものを選び出すのは、本研究とは、別の研究課題となる*。

4.2.2 MOVE層

働きかけや応答などの観点から見ると、ターン内の連続する複数のACT単位がまとまって1つの機能を果たす場合がある。MOVE層は、このようなターン内の同一の機能を持つ一連のACT単位をひとまとまりの単位(MOVE単位)として統合し、表現する。この単位が持つやりとり上の機能を、[開始],[応答],[補足]の3つのいずれかで記述する。ここで、[開始]は、質問や依頼のような聞き手に働きかける機能を、[応

答]は、質問への返答などの[開始]機能を持つ単位への応答を、[補足]は、[応答]された話者(すなわち[開始]発話を行った話者)がその応答への補足説明や評価を行う機能をそれぞれ表す。

発話表現を対応付けるうえで、対応付けるものどうしの個数が不定では扱いにくい。我々は、一度、ACT単位に分割した後、働きかけや応答などの観点から見て同じ役割を果たす発話をまとめることによりこの問題を回避している。たとえば、例1では、(1)で1個のACT単位で実現している機能を(2)では、2個で行っている。

(例1)

(1) 「お名前とご住所をお聞かせ願えますか？」

(2) 「お名前をいただけますか？」

「ご住所もお聞かせ願えますか？」

例2は、例1への応答の例である。ここでも同様に、(1)で1個のACT単位で実現している機能を(2)では、2個で行っている。

(例2)

(1) 「大阪の鈴木です」

(2) 「鈴木です」

「大阪に住んでいます」

このように、実際の発話では、同一の内容を様々な個数のACT単位で表現できる。しかし、(例1)(2)の発話「お名前をいただけますか？」と「ご住所もお聞かせ願えますか？」を統合し、1つの単位として扱うことで(1)の発話「お名前とご住所をお聞かせ願えますか？」と同様に扱うことができる。これは、(例2)(2)の発話についても同様である。このように、複数のACT単位を1つの単位に統合したMOVE単位をおくことで、発話を構成するACT単位の個数にかかわらず、発話表現間の対応関係をとることができるようになる。

4.2.3 EXCHANGE層

EXCHANGE層は、ある内容に関して、異なる話者が情報を交換する単位(EXCHANGE)を表現する。このEXCHANGEは、MOVE単位を構成要素として、対応関係を持った2個以上のMOVE単位から構成される。また、対応するMOVE単位は、その機能の組合せとして下記の2つのパターンのいずれかとなる。

(1) [開始] + [応答]

(2) [開始] + [応答] + [補足]

実際の対話では、「[開始1][開始2]…[応答2][応答1][補足1]」のように入れ子状に現れうる。この場合、[開始1]と[応答1][補足1]、[開始2]と[応

* 発話の対応関係は、正しい発話行為を選択するのにも利用できる場合がある。たとえば、「予約をお願いします」という発話は、対話開始時に用いられた場合は、依頼と解釈できるが、「市内観光バスの予約はなさいますか」という発話の後に用いられた場合は、質問への応答と解釈できる。

表5 対話例
Table 5 A dialogue example.

UID	話者	発話
...
U7	S	登録用紙はすでにお持ちでしょうか
U8	Q	いいえ
U9	Q	まだです
U10	S	分かりました
U11	S	それではこちらからお送りいたしますので
U12	S	お名前とご住所をお聞かせ願えますか
U13	Q	はい
U14	Q	大阪市北区茶屋町六の二十三鈴木真弓です
U15	S	はい
U16	S	分かりました
U17	S	では至急に送らせていただきます
U18	Q	参加料はいるのでしょうか
U19	S	はい
U20	S	登録費としてお1人一万六千円です
...

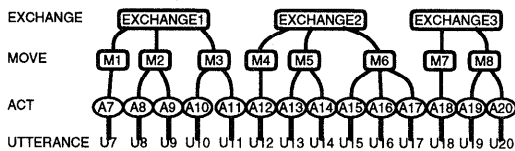


図1 対話構造の例

Fig. 1 An example of a dialogue structure.

答2] がそれぞれ対応関係にある。ACT層、MOVE層を介してEXCHANGE層で発話間の対応関係を表現することで、上記のような簡潔な形式で関係を表現できる。

以下、[開始]機能と[応答]機能とをそれぞれ持つ2個のMOVE単位から構成されるEXCHANGEを「成立したEXCHANGE」と呼ぶ。

4.2.4 対話構造例

前述した対話構造のモデルに従うと、表5に示す対話例全体は、図1に示す対話構造として表現される。図1中において、各発話を表5中のUIDを用いて表現している。

たとえば、表5の対話例の発話U18~U20では質問-応答のやりとりが行われている。この場合の対話構造は次のようになる。発話U18は、質問にかかわるACT単位としてACT層に記述される。このACT単位が構成するMOVE単位の持つ機能は[開始]であり、1個のMOVE単位「M7」を形成する。また、同様に発話U19は、肯定についてのACT単位として、発話U20は、情報伝達のACT単位としてACT層に記述される。これら2つのACT単位は、MOVE層上の機能として1つの[応答]を持ち、1個のMOVE単位「M8」を形成する。さらにこれら2個の[開始]と[応答]の

MOVE単位が1個のEXCHANGE「EXCHANGE3」を形成している。このようにして発話U18、U19、U20の対応関係が表現される。

4.3 対話構造認識手法

対話の展開とともに、発話をACT単位に分割し、各単位をMOVE単位、EXCHANGEとまとめあげていくことで、対話構造を漸次的に形成する。本手法は、発話ごとに次のステップに従って処理が進行する。
ステップ0：発話行為を行いうる最小単位（ACT単位）への分割

ステップ1：働きかけ、応答の観点から見た機能への分類（ACT単位の表現）

ステップ2：働きかけ、応答の観点から見たターン中の最小単位への統合（MOVEの形成）

ステップ3：働きかけ、応答の構造の生成（EXCHANGEの形成）

上記ステップ0は、前処理であり、分割の結果できる各ACT単位ごとにステップ1~3を繰り返す。対話構造を認識するうえで生じる働きかけ、応答の観点から見た発話の持つ機能の多義は、ステップが進むにつれて解消される。以下、上記の順に従って各ステップについて述べる。

4.3.1 ステップ0：発話行為を行いうる最小単位（ACT単位）への分割

まず、構造を求める前に、入力発話表現を発話行為を行いうる最小単位（ACT単位）へ分割する。分割は、分割後の各単位がただ1つの発話行為を持つように行う。この分割は、あらかじめ定義した表層表現パターンを用いて行う。たとえば、発話「それではこちらからお送りいたしますのでお名前とご住所をお聞かせ願えますか」は、どちらの単位も単独で現れても発話として成立するので、「それではこちらからお送りいたしますので」と「お名前とご住所をお聞かせ願えますか」の2つのACT単位へ分割される。

4.3.2 ステップ1：働きかけ、応答の観点から見た機能への分類（ACT単位の表現）

ステップ1では、ACT単位を働きかけ、応答の観点から見た機能に分類する。この段階では、分類しうるすべての機能を列挙する。これは、

- (1) 発話行為の認定
- (2) 働きかけ、応答の観点から見た発話行為の機能への分類

の2段階で行われる。

4.3.2.1 発話行為の認定

ACT単位ごとに発話行為を認定する。発話行為は、あらかじめ定義した表層表現パターンと発話行為の対

表6 ACT 辞書 (一部)
Table 6 ACT dictionary.

表層表現パターン	ACT ラベル
はい	<greet>, <acknowledge>, <yes>
*です	<inform>
すいません	<thank>, <apology>, <alert>
たいんです	<desire>
*どのような*か	<wh-question>
...	...

応を示すテーブルを用いて認定し、25種類のACTラベル*を用いて表現する。発話行為に関して、多義が生じた場合は、すべてを候補として列挙する。

上記のテーブルをACT辞書と呼び、内容の一部を表6に示す。この表において、表層表現パターン中の「*」は、表現中のその位置に任意の表現を許すことを示している。また、複数のACTラベルが記述されているものは、行為の多義があることを示す。

表層表現パターンと発話の照合の際、複数のパターンと一致した場合、最長のパターンが採用される。たとえば、「どのようにお支払いしたらよいのですか」という発話に、「*どのように*か」と「*ですか」の2個の表層表現パターンが一致した場合、前者のパターンが採用され、ACTラベル<wh-question>が付与される。

ACT辞書は、人手で対話データを参照することにより作成する。辞書作成者が、発話を持つ発話行為を認定するのに必要な文末表現や手がかり語の表層表現パターンを抽出し、このパターンと、発話行為を対応させて記録する。1つのパターンが複数の発話行為と対応する場合は、すべての発話行為を網羅的に記録する。したがって、パターン中には、発話行為に関する多義を持つものがある。

ここで用いる表層表現パターンは、領域知識に依存しない助詞、助動詞、接続詞等の組合せで構成される文末表現や手がかり語である。このため、同一のACT辞書を様々な領域に適用することが可能である。

4.3.2.2 働きかけ、応答の観点から見た発話行為の機能への分類

このステップでは、働きかけ、応答の観点から見た機能へACT単位を分類する。分類は、ACT単位が持つ発話行為と機能の対応関係を記述したACT-機能対応テーブル(表7)を参照して行う。

ACT単位には、各単位単独で機能が一意に決定できるものと、対話構造の状態を参照する必要のあるも

表7 ACT-機能対応テーブル
Table 7 ACT-function table.

ACT ラベル	機能
<accept>	[応答],[補足]
<acknowledge>	[応答],[補足]
<action-request>	[開始]
<alert>	[開始]
<apology>	[開始]
<believe>	[開始],[応答],[補足]
<confirmation-question>	[開始]
<desire>	[開始],[応答],[補足]
<do-you-understand-question>	[開始]
<expressive>	[開始],[応答],[補足]
<farewell>	[開始],[応答],[補足]
<good-wishes>	[開始],[応答],[補足]
<greet>	[開始],[応答],[補足]
<inform>	[開始],[応答],[補足]
<instruct>	[開始]
<no>	[応答],[補足]
<offer-follow-up>	[開始],[応答],[補足]
<permission-request>	[開始]
<promise>	[開始],[応答],[補足]
<temporizer>	[開始],[応答],[補足]
<thank>	[開始]
<thanks-response>	[応答],[補足]
<wh-question>	[開始]
<yes>	[応答],[補足]
<yn-question>	[開始]

のがあ。たとえば、発話「参加なさいますか」は、聞き手へ働きかける発話であり、単独で機能を[開始]と決定できる。しかし、発話「コーヒーが飲みたいのですが」の機能を決定するためには、先行発話との関係、つまり、対話構造の状態を参照する必要がある。この表現が、コーヒーを要求する場面で発話されたのであれば、やりとり上の機能は[開始]となる。しかし、「何かお飲みになりますか」に対応するものとして発話されたのであれば、[応答]となる。さらに、[補足]ともなりうる。

このステップでは、まだ対話構造の状態を参照していないので上記のように機能を一意に決定できない場合が生じる。この機能の多義をその生じ方に応じて以下の3つに分類する。

機能クラス1:[開始]

機能クラス2:[応答],[補足]

機能クラス3:[開始],[応答],[補足]

4.3.3 ステップ2:働きかけ、応答の観点から見た最小単位への統合(MOVE単位の形成)

本ステップでは、同一ターン内にあり、同一の機能クラスに分類された連続して生起する1個以上のACT単位を単一の単位(MOVE単位)に統合する。

まず、ACT単位が持つ機能の多義を以下の機能クラス解釈ルールに従って、機能クラス1または機能ク

☆ここで使用している分類は、Seligmanら¹⁶⁾が提案したものである。

ラス 2 のいずれかに解釈する。

機能クラス解釈ルール

- ACT 単位が機能クラス 3 で次のいずれかを満たす場合は機能クラス 1 とする, それ以外の場合は機能クラス 2 とする
 - － 対話開始時の第一 ACT 単位である
 - － 話者交替直後の第一 ACT 単位でかつすべての EXCHANGE が成立している
 - － 直前 ACT 単位の話者と同一話者であり, 直前 ACT 単位が機能クラス 1 に分類される
 - 機能クラス 1 また 2 である場合はそのままとする
- MOVE 単位の生成, 統合は, 以下に示す MOVE 生成ルールに従って行う。

MOVE 生成ルール

- 次のいずれかを満たす ACT 単位は, 新たな MOVE 単位を生成する
 - － 対話開始時の第 1 ACT 単位
 - － 話者交替直後の第 1 ACT 単位
 - － 直前 ACT 単位と同一のターン内にあり, 機能クラスが直前の ACT 単位と異なる場合
- 新たな MOVE 単位を生成する場合, 直前の ACT 単位で MOVE 単位は終了する
- 直前 ACT 単位と同一話者でかつ同一機能クラスである場合, 直前 ACT 単位の属する MOVE 単位が継続する

たとえば, 話者交替が生起したり, 直前の ACT 単位の機能クラスと比較して, 機能クラス 1 → 機能クラス 2, あるいは, 機能クラス 2 → 機能クラス 1 のように機能クラスが変化した場合, すでにある MOVE 単位が終了し, 新たな MOVE 単位が生成される。さもなければ, 直前の ACT 単位が属している MOVE 単位が継続していると解釈され, それに統合される。

4.3.4 ステップ 3 : EXCHANGE の形成

このステップでは, 前ステップで作成された MOVE 単位間の対応関係を決定し, 働きかけ, 応答の単位である EXCHANGE を形成する。このとき, 新たな EXCHANGE は, 機能クラス 1 を持つ新たな MOVE 単位の生成とともに生成され, 機能クラス 2 を持つ MOVE 単位の終了とともに終了される。つまり, [開始] + [応答] または [開始] + [応答] + [補足] の組合せが成立した後, 終了する。EXCHANGE の組合せが成立した時点で, 発話間の対応関係が認識されたことになる。

新たな MOVE 単位が生成されると, MOVE 単位の機能を以下に示す MOVE 機能決定ルールに従って決定する。ここで, 各機能を持つ MOVE 単位をそれ

ぞれ [開始] MOVE, [応答] MOVE, [補足] MOVE と略記する。

MOVE 機能決定ルール

- 機能クラス 1 は, [開始] とする (定義より)
- 次の場合, 機能クラス 2 は, [応答] とする
 - － 構成要素として [応答] MOVE を持たない EXCHANGE が存在する場合☆
- 次の場合, 機能クラス 2 は, [補足] とする
 - － すべての EXCHANGE がその構成要素として [開始] MOVE と [応答] MOVE の両方を持つ場合**

このように, 機能クラス 2 に分類されていた MOVE 単位の持つ機能の多義が構成要素として [応答] MOVE を持つ EXCHANGE の存在の有無***により解消される。すなわち, 機能クラス 2 は, 対応する [応答] MOVE が決まっていな [開始] MOVE がある場合は [応答] に, ない場合は [補足] と解釈される。

次に, 以下に示す EXCHANGE 形成ルールに従って EXCHANGE を形成する。

EXCHANGE 形成ルール

- [開始] MOVE が新たに生成されたとき, 新たな EXCHANGE が生成される
- [応答] MOVE が新たに生成されたとき, 構成要素として [応答] MOVE を持たない最も新しく生成された EXCHANGE の構成要素とする
- [応答] MOVE が終了したとき, EXCHANGE は終了する
- MOVE 単位が継続している場合は, EXCHANGE は継続する
- 先行するすべての EXCHANGE が終了しており, [補足] MOVE が新たに生成された場合, 終了した EXCHANGE で最新のものにその構成要素として繰り込まれる。

4.4 処理例

表 5 中の一連の ACT 単位 U12~U17 を例に, その機能の多義解消と対応付けを行った過程の概略を表 8 に示す。

表 8 において, ステップ 1 欄の ACT ラベル欄は, ACT 辞書を用いて認定された発話行為を, 機能欄は,

☆ 対応する [応答] MOVE が決定されていない [開始] MOVE が存在する場合である。

** 先行するすべての [開始] MOVE の各々に対応する [応答] MOVE が 1 個ずつ存在する場合である。

*** 対応する [応答] MOVE の未決定の [開始] MOVE の有無を示す。

表8 対話構造認識処理過程

Table 8 Process for recognizing a dialogue structure.

UID	話者	発話	ステップ1			ステップ2		ステップ3	
			ACT ラベル	機能	機能クラス	機能クラス	MOVE	機能	EXCHANGE
...
U11	S	それではこちらからお送りいたしますので	<inform>	[開始] [応答] [補足]	3	2	継続/終了	[補足]	繰込/終了
U12	S	お名前とご住所をお聞かせ願えますか	<action-request>	[開始]	1	1	生成/終了	[開始]	生成
U13	Q	はい	<greet> <acknowledge> <yes>	[応答] [補足]	2	2	生成	[応答]	継続/終了
U14	Q	大阪市北区茶屋町六の二十三鈴木真弓です	<inform>	[開始] [応答] [補足]	3	2	継続/終了		
U15	S	はい	<greet> <acknowledge> <yes>	[応答] [補足]	2	2	生成		
U16	S	分かりました	<inform>	[応答] [補足]	2	2	継続	[補足]	継続/繰込
U17	S	では至急に送らせていただきます	<promise> <inform>	[開始] [応答] [補足]	3	2	継続/終了		
U18	Q	参加料はいるのでしょうか	<wh-question> <yn-question>	[開始]	1	1	生成	[開始]	生成
...

ACT-機能対応テーブルを参照して ACT 単位が分類された機能を、機能クラス欄は、機能の多義の生じ方に応じて分類された機能クラス名をそれぞれ示す。また、ステップ2欄の機能クラス欄は、機能クラス解釈ルールの適用結果を、MOVE欄は、MOVE生成ルール適用結果をそれぞれ示す。このMOVE欄中で、「生成」は、新たなMOVE単位が生成されることを、「継続」は、直前のACT単位が属しているMOVE単位が継続しており、その時点でのACT単位がそのMOVE単位に統合されることを、「終了」は、次のACT単位が産出された時点で、そのMOVE単位が終了することを、それぞれ示す。さらに、ステップ3欄の機能欄は、MOVE機能決定ルールの適用結果を、EXCHANGE欄は、EXCHANGE形成ルールの適用結果をそれぞれ示す。このEXCHANGE欄中で、「生成」は、新たなEXCHANGE単位が生成されることを、「継続」は、直前のMOVE単位が属しているEXCHANGEが継続しておりその時点でのMOVE単位がそのEXCHANGEに統合されることを、「終了」は、次のMOVE単位が産出された時点で、そのEXCHANGEが終了することを、「繰込」は、その時点でのMOVE単位が、直前に終了したEXCHANGEに繰り込まれることを、それぞれ示す。

たとえば、ACT単位U12:Q:「大阪市北区茶屋町六の二十三鈴木真弓です」は、以下のように処理される。

ステップ1: ACT辞書を用いてACTラベル(inform)が付与される。さらに、ACT-機能対応テーブルから機能[開始],[応答],[補足]に分類される。この結果、機能の多義の生じ方から機能クラス3となる。

ステップ2: まず、機能クラス解釈ルールを適用する。この結果、機能クラス2と解釈される。次に、MOVE生成ルールを適用する。このACT単位は、直前のACT単位U15「はい」の話者Qと同一であり、機能クラス2である。この結果、直前のACT単位「はい」が属するMOVE単位が継続し、そこへ統合される。

ステップ3: MOVE単位の機能は、直前ACT単位の属するMOVE単位が継続していることからその機能[応答]が継続する。この機能は、ACT単位U13「はい」により新たなMOVE単位が生成された時点で、MOVE機能決定ルールを適用することにより決定される。ACT単位U13の機能クラスは2であり、ACT単位U12の時点で生成された構成要素として[応答]MOVEを持たな

い EXCHANGE が存在する。この結果、上記のように機能が「応答」であるとされている。次に EXCHANGE 形成ルールを適用する。MOVE 単位が継続していることから、すでに存在する EXCHANGE が継続する。この EXCHANGE は、ACT 単位 U12 からなる「開始」MOVE と ACT 単位 U13 と U14 とからなる「応答」MOVE を構成要素としている。この結果、この時点で ACT 単位 U12, U13, U14 の間に対応関係があることが認識できる。

5. 評価実験

本研究の目的は、対話構造を通して対話中の発話表現間の対応関係を認識することである。この対応関係の認識精度を実験により測定し評価した。認識精度は、本手法により得られた発話表現間の対応関係と人間が認定した対応関係とを比較することによって測定した。

5.1 実験対象対話データ

評価対象とした対話データは、「旅行」に関する 200 対話（書き起こし文字数^{*}: 126,083 文字、ターン数: 3,781 個）である。同一の対話データに対して、6 人の被験者が発話間の対応関係を付与した。各被験者が 200 対話すべてに対して作業を行った。被験者に提示した対話は、すべて人間同士の対話を収録し書き起こしたもので、各ターンごとに話者の違いが明示されている。被験者には、質問-応答のような対応関係にある発話表現に同一の ID を付与するように指示した。また、同一ターン内に対応関係が異なる表現が含まれている場合には、そのターンを対応関係に従って分割するよう指示した。以下、便宜的に分割後の各表現を発話単位と、また、同一 ID が付与されている発話単位全体を交換単位と略記する。たとえば、表 9 の例では、話者 B の発話「はいございますお召し上がりになりますか」が「はいございます」と「お召し上がりになりますか」の 2 つの発話単位に分割され、さらに、「チーズケーキはありますか」と「はいございます」に同一の ID が、「お召し上がりになりますか」と「1つお願いします」に前出のものとは異なる同一の ID が付与される。ただし、発話の分割、対応付けは、被験者の判断に任せたので被験者間で一致するとは限らない。

5.2 実験方法

実験対象対話データに対して本手法を適用し、本手法と被験者が付与した発話間の対応関係を比較する

表 9 作業指示例

Table 9 A dialogue example in the instruction for the human subjects.

話者	発話
A:	チーズケーキはありますか
B:	はいございますお召し上がりになりますか
A:	1つお願いします

ことによって認識精度を測定した。

5.2.1 準備

まず、実験対象対話データ 200 対話からクローズドセットとして 50 対話をランダムに抽出し、残り 150 対話をオープンセットとした。次にクローズドセットを用いて ACT 辞書を作成した。1 人の作業者が、クローズドセット中の発話を持つ発話行為とそれを認定するのに必要な文末表現や手がかり語の表層表現パターンを抽出し、このパターンと発話行為を対応させて記録した。ここで提示する実験では、すべてここで作成した ACT 辞書を使用している。

5.2.2 調査項目

次の 3 項目について調査した。

- (1) 被験者同士に対応関係付けの一致の精度
- (2) 本手法による対応関係付け結果と被験者による結果の一致の精度
- (3) 単純な手法による対応関係付け結果と被験者による結果の一致の精度

項目 (1) は、被験者による結果の信頼性、一貫性を確認するための調査である。また、項目 (2) は、本手法の有効性を確認するための調査である。さらに、項目 (3) は、本手法の効果を確認するための対照調査である。

単純な手法として次の 2 種類を用いた。

- (1) 1 個のターンが単一の MOVE 単位を構成し、隣接する 2 個のターンが EXCHANGE を構成すると見なす手法
- (2) 1 個のターンが単一の MOVE 単位を構成し、隣接する 3 個のターンが EXCHANGE を構成すると見なす手法

すなわち、単純な手法 (1) は、対話を構成する EXCHANGE がすべて「開始」MOVE + 「応答」MOVE の構成を、単純な手法 (2) は、「開始」MOVE + 「応答」MOVE + 「補足」MOVE の構成をそれぞれしていることを想定している。また、入れ子状の構造は、まったく想定していない手法である。

5.2.3 評価尺度

前項で述べた被験者に対する指示より交換単位に含まれる発話単位は相互に対応関係にある。また、被

^{*} 漢字かな混じり表現として書き起こしたものである。

験者による発話単位, 交換単位は, 本手法における, ACT 単位, EXCHANGE にそれぞれ対応する. 認識精度は, ACT 単位ごとに, その ACT 単位が属する EXCHANGE と対応する交換単位について, 各々に属する ACT 単位と発話単位の一一致数を用いて測定する.

認識精度の尺度として適合率, 再現率を用いた. 再現率, 適合率の定義を次に示す.

$$\text{再現率} = \frac{\text{一致 ACT 単位数}}{\text{交換単位中の発話単位数}} \quad (1)$$

$$\text{適合率} = \frac{\text{一致 ACT 単位数}}{\text{EXCHANGE 中の ACT 単位数}} \quad (2)$$

実験では, 各対話中のすべての ACT 単位について上記の再現率, 適合率を算出し, 対話ごとに平均することで, 対応関係の認識精度を測定した.

5.3 実験結果

被験者による発話表現間の対応関係付け結果と本手法による結果を比較することにより本手法の有効性を評価した.

まず, 被験者同士に対応関係付けの一致の精度の調査を行い, 被験者による結果の信頼性, 一貫性を確認した. 調査は, 6 人中のすべての 2 被験者の組合せごとに精度を計算することで行った. その結果を表 10 に示す. 被験者同士の適合率, 再現率が, クローズドセット, オープンセットのどちらに対しても約 82% であることから, 評価対象として, 信頼性, 一貫性のあるデータであり, 本手法の評価対象として有効であることが確認できた.

次に, 本手法による対応関係付け結果と被験者による結果とを比較し, 精度を調査した. 表 11 に, 6 人の被験者 A~F 各々に対する本手法の精度を示す. 6 人の被験者に対する適合率, 再現率の平均は, クローズドセットに対して, それぞれ 72.9%, 80.6% であり, オープンセットに対しては, それぞれ 71.8%, 78.1% であった. また, 対象データをクローズドセットからオープンセットに変更した場合の精度の低下は, 3% 未満にとどまっている. これらのことから, 本手法が, 被験者に近い精度で対応関係を認識でき, 適用領域に対して一般性があり, オープンデータに対して頑健であることが確認できた.

また, 表 12 に, 6 人の被験者すべてが共通に対応付けをした発話単位に関する精度 (表 12 中では「すべて」と表記) と被験者のうち少なくとも 1 人以上が対応付けをした発話単位に関する精度 (表 12 中では「最低 1 人」と表記) を示す. 被験者「すべて」が認識した対応関係は, 信頼性が高い関係であると考えられる. この関係に対して, 再現率が, クローズドセ

表 10 被験者同士の発話間の対応関係認識の一致度

Table 10 The agreement accuracies between human subjects' judgements.

	クローズド		オープン	
	適合率	再現率	適合率	再現率
被験者同士	82.1%	82.1%	81.9%	81.9%

表 11 被験者と本手法の発話間の対応関係認識の一致度 (1)

Table 11 The agreement accuracies between the recognition results and human subjects' judgements (1).

被験者	クローズド		オープン	
	適合率	再現率	適合率	再現率
A	79.7%	80.6%	79.5%	76.0%
B	67.3%	87.4%	66.9%	85.7%
C	75.3%	73.2%	71.6%	68.6%
D	69.2%	84.7%	64.2%	83.5%
E	73.0%	74.9%	76.6%	74.1%
F	73.2%	82.5%	71.7%	80.9%
平均	72.9%	80.6%	71.8%	78.1%

表 12 被験者と本手法の発話間の対応関係認識の一致度 (2)

Table 12 The agreement accuracies between the recognition results and human subjects' judgements (2).

手法	被験者	クローズド		オープン	
		適合率	再現率	適合率	再現率
本手法	すべて	53.6%	94.2%	53.9%	93.1%
本手法	最低 1 人	89.1%	64.3%	87.4%	60.1%

表 13 被験者と単純な手法による発話間の対応関係認識の一致度

Table 13 The agreement accuracies between the recognition results with simple methods and human subjects' judgements.

手法	クローズド		オープン	
	適合率	再現率	適合率	再現率
単純な手法 (1)	60.0%	72.1%	61.2%	70.5%
単純な手法 (2)	46.9%	80.5%	48.4%	79.3%

トで, 94.2%, オープンセットで, 93.1% を得ていることから, 本手法は, 信頼性の高い関係を高精度で認識できることが分かる. また, 被験者「最低 1 人」が認識した対応関係は, 少なくとも 1 人の被験者が対応関係を認めており, 対応を持つ可能性がある関係と考えられる. この関係に対して, 適合率が, クローズドセットで 89.1%, オープンセットで 87.4% を得ている. これは, 本手法が認識した対応関係のうち, 被験者のだれ 1 人も対応関係を認めなかった関係 (非対応関係) の割合が, 約 10% であることを示しており, 誤認識が少ないことを示している.

さらに, 表 13 に, 単純な手法 (1), (2) による対応関係付け結果と被験者による結果の一致の精度を示

す^{★1}。単純な手法(1)による適合率, 再現率は, それぞれ約60%, 約70%であり, 本手法が, 適合率, 再現率ともに約10%上回っている。また, 単純な手法(2)による適合率, 再現率は, それぞれ約47%, 約80%である。したがって, 本手法が, 再現率は同等, 適合率が約25%上回っている。これは, 単純な手法(2)による認識結果には, 誤った対応関係が多く含まれていることを示している^{★2}。これらのことから, 固定個数の隣接するターンを対応させてEXCHANGEを形成するのでは, 適切な対応関係をとらえることができないことが確認できた。これは, MOVE単位を導入し, 隣接関係だけでなく入れ子関係を考慮して対応付けし, さらに, 対応付けるMOVE単位の個数は, 状況に応じて変化させている本手法が, 適切な対応関係をとらえるのに効果があることを示している。

実験結果をまとめると下記のようになる。

- 被験者による対応付けは, 一貫性があり, 信頼性がある
- 人間に近い精度で対応関係を認識できる
- 一般性があり, オープンデータに対して頑健である
- 信頼性が高い対応関係^{★3}を高い精度で認識する
- 非対応関係^{★4}に対する誤認識は少ない

これは, 本手法が発話表現間の対応付けに関して有効であることを示している。

5.4 実験結果の考察

本手法では, パターン「[[開始] + [応答]]」または「[[開始] + [応答] + [補足]]」のいずれかに従って発話表現を対応付ける。したがって, 機能が「開始」であると認定された発話表現には, 後続する発話表現のいずれかを必ず対応付ける。このため, 対話開始時に現れる「あいさつ」のように対応する発話表現がない場合は, 誤った対応関係を認識してしまう場合がある。また, 発話の命題内容に関する情報を使用していないため, 表14に示すような, 2つの異なるターンでなされた質問(発話番号:1, 3)に1回の発話(発話番号:4)で応答をしてしまう場合も誤った対応関係を認識してしまう。しかし, 6人の被験者のうちいずれか1人が対応関係を認定した発話単位に対する適

表14 対応付けができない例
Table 14 A dialogue example failing to recognize relations.

発話番号	発話者	発話表現
1	駅員	ファーストとコーチどちらがよろしいでしょうか
2	お客	お値段はそれぞれいくらくらい違うんですか
3	駅員	片道ですかそれとも往復をお買いになりますか
4	お客	片道で結構ですファーストにしてください

合率が約90%であり, 誤った対応関係を認識してしまうことは, 少数であることから, 実験対象データ中では, 上記のような例は, 少ないと考えられる。しかし, これが, 対話の持つ一般的な性質であるか否かの確認は, 今後の研究課題である。

6. まとめ

本論文では, 省略補完や代用表現の解釈に有用な発話表現間の対応関係を得ることを目的として, 対話構造のモデルを提示し, このモデルに基づいた対話構造認識手法を提案した。このモデルでは, ACT, MOVE, EXCHANGEの3階層の構造で対話を表現する。

ACT層では, 発話を単一の発話行為を持つ発話単位に分割し, 各発話単位を発話行為を用いて表現する。MOVE層では, 複数の発話単位を働きかけ-応答の機能の観点から1つの単位にまとめ, 表現する。さらに, EXCHANGE層では, MOVE層上の単位を用いて発話間の対応関係を表現する。

対話構造の認識は, 発話ごとに, 発話表現の分割, 発話行為への分類, 機能への分類, MOVE単位へのまとめあげ, EXCHANGEの形成のステップで行われる。この過程で, 本論文で提案した対話構造認識手法は, 対話のかかわる話題領域の知識を用いていない。このため領域の違いにかかわらず同様に動作する。したがって, 本稿で実験対象データとして用いた「旅行会話」だけでなく, 他の領域にかかわる対話にも適用可能である。

本手法の有効性を確認するために, 実験により発話表現間の対応付けを行い, (1)被験者間, (2)本手法-被験者間の対応付け結果を定量的に比較, 評価した。また, 単純な手法を用いて対照実験をし, 単純な手法-被験者間の対応付け結果を定量的に比較することで本手法の効果を確認した。

被験者間の対応付けの比較結果は, 適合率, 再現率がどちらも約82%であり, 被験者による対応付け結

^{★1} 6人の被験者に対する適合率, 再現率の平均のみを示す。

^{★2} これは, この手法では, 1個のEXCHANGEは, 隣接する3個のターンから形成されており, 多くの発話単位から構成されているためである。同様に対話全体を1個のEXCHANGEと見なせば, 再現率100%が得られる。このとき, 当然, 適合率は, 低下する。

^{★3} 被験者全員が認めた対応関係。

^{★4} 被験者全員が認めなかった対応関係。

果は、信頼性、一貫性があり、本手法の評価対象として有効であることが確認できた。評価実験により得られた本手法-被験者間での対応付け比較結果は、オープンデータに対して、平均で、適合率71.8%、再現率78.1%である。さらに、これらは、クローズドデータの結果に対して3%未満の低下にとどまっている。また、同様にオープンデータに対して、被験者6人が対応関係を認めた発話単位に対する本手法の対応関係の再現率は、93.1%であり、被験者の少なくとも1人以上が対応関係を認めた発話単位に対する本手法の対応関係の適合率は、87.4%であった。これは、本手法が、適用領域にかかわらず、人間に近い精度で発話の対応関係を認識でき、特に、信頼性が高い対応関係を高い精度で認識し、非対応関係に対する誤認識は少ないことを示している。以上より、本手法が、発話間の対応関係の認識の目的に有効であることが確認できた。

謝辞 本研究は、著者3人全員が、ATR 音声翻訳通信研究所に在籍中に行ったものです。本研究の機会を与えてくださるとともに適切な助言をくださった当時の同研究所山崎泰弘社長に感謝します。また、熱心に討論して下さった同研究所第四研究室の諸氏に感謝します。貴重なコメントをいただいたNTTコミュニケーション科学研究所加藤恒昭主幹研究員に感謝します。

参考文献

- Allen, J.F.: Recognizing intentions from natural language utterances, *Computational Models of Discourse*, Brady, M. and Berwick, R.C. (Eds.), pp.107-166, MIT Press, Cambridge, MA (1983).
- Grosz, B.J. and Sider, C.L.: Attention, intention and the structure of discourse, *Computational Linguistics*, Vol.12, No.3, pp.175-204 (1986).
- Grosz, B. and Hirschberg, J.: Some International Characteristics of Discourse Structure, *Proc. International Conference on Spoken Language Processing (ICSLP)* (1992).
- Hirschberg, J. and Grosz, B.J.: Intonational features of local and global discourse structure, *Proc. Darpa Workshop on Spoken Language* (1992).
- 飯田 仁, 有田英一: 4階層プラン認識モデルを使った対話の理解, 情報処理学会論文誌, Vol.31, No.6, pp.810-821 (1990).
- 石崎雅人, 小磯花絵: 対話研究の新しい流れ, SIG-SLUD 9503-2, 人工知能学会 (1995).
- J., P.R. and Litman, D.J.: Intention-Based Segmentation: Human Reliability and Correlation with Linguistic Cues, *Proc. 31st ACL* (1993).
- Katoh, N. and Morimoto, T.: Statistical Method of Recognizing Local Cohesion in Spoken Dialogues, *Proc. COLING-96* (1996).
- Kay, M., Gawron, J.M. and Norvig, P.: Verbmobil: A Translation System for Face-to-Face Dialog, CSLI Lecture Notes 33, CSLI (1994).
- 熊本忠彦, 伊藤 昭: 支援対話解析に基づく対話処理方式, 電子情報通信学会論文, Vol.J77-D-II, No.8, pp.1492-1501 (1994).
- 熊本忠彦, 伊藤 昭, 海老名毅: 支援対話におけるユーザ発話意図の認識—ユーザ発話文の解析に基づく統計的アプローチ, 電子情報通信学会論文, Vol.J77-D-II, No.6, pp.1114-1123 (1994).
- Litman, D.J. and Allen, J.F.: A Plan Recognition Model for Subdialogues in Conversations, *Cognitive Science*, Vol.11, No.2, pp.163-200 (1987).
- Litman, D.J. and J., P.R.: Combining Multiple Knowledge Sources for Discourse Segmentation, *Proc. 33rd ACL* (1995).
- Nagata, M. and Morimoto, T.: An Information - Theoretic Model of Discourse for Next Utterance Type Prediction, *Trans. Information Processing Society of Japan*, Vol.35, No.6, pp.1050-1061 (1994).
- Reithinger, N. and Maier, E.: Utilizing Statistical Dialogue Act Processing in Verbmobil, *Proc. 33rd ACL* (1995).
- Seligman, M., Fais, L. and Tomokiyo, M.: A Bilingual Set of Communicative Act Labels for Spontaneous Dialogues, ATR テクニカルレポート, TR-IT-0081, ATR 音声翻訳通信研究所 (1995).
- Stenstrom, A.-B.: *An Introduction to Spoken Interaction*, Longman (1994).
- Waibel, A., Jain, A., McNair, A., Saito, H., Hauptmann, A. and Tebelskis, J.: JANUS: A Speech-to-Speech Translation System Using Connectionist and Symbolic Processing Strategies, *Proc. ICASSP '91*, pp.793-796 (1991).
- 久米雅子, 小暮 潔: 発話意図の翻訳のための発話行為推論部, 第39回情報処理学会全国大会論文集, 3F-5, pp.612-613 (1989).
- 森元 逞, 田代敏久, 竹澤寿幸, 永田昌明, 谷戸文廣, 浦谷則好, 鈴木雅実, 菊井玄一郎: 音声翻訳実験システム (ASURA) のシステム構成と性能評価, 情報処理学会論文誌, Vol.37, No.9 (1996).

(平成9年11月12日受付)

(平成10年6月5日採録)

**巖寺 俊哲 (正会員)**

1963年生。1986年電気通信大学情報数理工学科卒業，1988年同大学大学院修士課程修了。同年日本電信電話株式会社に入社。1993年から1996年までATR音声翻訳通信研究所に出向。現在，NTT情報通信研究所に所属。情報検索，自然言語処理の研究・開発に従事。

**石崎 雅人 (正会員)**

1960年生。1984年慶應義塾大学大学院修士課程修了。同年日本電信電話株式会社入社。1994年から1997年までATR音声翻訳通信研究所に出向。1998年北陸先端科学技術大学院大学助教授，現在に至る。ph.D. (英エディンバラ大学認知科学科・人工知能学科)。人工知能学会，言語処理学会，国際計算言語学会各会員。

**森元 暎 (正会員)**

1946年生。1970年九州大学大学院修士課程修了。同年日本電信電話公社入社。1987年よりATR自動翻訳電話研究所ならびにATR音声翻訳通信研究所に勤務。1998年福岡大学電子情報工学科教授，現在に至る。工学博士。電子情報通信学会，人工知能学会，言語処理学会，日本音響学会各会員。