

料理教示発話の構造解析

西田悠介 柴田知秀 河原大輔 岡本雅史 黒橋禎夫 西田豊明

東京大学 大学院情報理工学系研究科

1 はじめに

計算機、ネットワーク、さらに種々のメディア処理技術の進展により、人間の知的作業に対する計算機の高度なサポートが可能となりつつある。たとえば、Kiyotaらのダイアログナビは、マイクトソフトのサポート技術情報などの自然言語テキストに基づき、MS製品に対するユーザの質問に対話的に応答する[1]。

このようなシステムの発展として、さまざまな動作・作業について、方法、注意点、コツなどを対話的・映像的に教示するシステムを考えることができる。工作、料理、電気製品の使い方、スポーツ習得など、これまで専門家、教師などに、直接、見本を示してもらいながら指導してもらうことが理想的であった分野について、このようなシステムは非常に有効である。

このようなシステムを実現するためには、専門家、教師などの実際の説明(映像)を構造化する必要がある。そのためにまず必要となるのは、教示発話を言語的に解釈することである。本論文では、このような位置付けの中で、料理教示発話を具体的題材として、その構造解析を行う。なお、ここでは音声認識は対象外とし、料理教示発話の書きおこしテキストを出発点とする。

2 料理教示発話の構造

2.1 料理教示発話のタイプ分類

料理教示発話は、実際に行う個別の作業の説明が中心であるが、その間に、ひとまとまり作業を始めることの宣言や、作業のコツ、注意事項などが述べられる。このような料理教示発話のタイプは次のように分類できる[2]。なお、ここではタイプ分類の単位は発話ごと(書きおこしテキストにおいて句点で区切られたもの)としている。複文などで一発話中に複数の単位を考え

る問題については2.3節で述べる。

作業宣言 (ひとまとまりの作業を始める宣言)

- ・ さ、では、ステーキの材料にかかります。
- ・ まず、キャベツです。

個別作業 (具体的な作業)

- ・ お鍋にお水を入れます。
 - ・ 食べやすい大きさにこうしてちぎってください。
- ※ 動作そのものが述語で、動作に関する食品、道具、時間、程度(強火で、たっぷり、など)、条件、理由などが項または従属節で示される。

料理状態

- ・ 高野豆腐と椎茸が、煮えております。

食品・道具提示

- ・ 材料は、牛ひき肉、百五十グラムです。
- ・ これを今日はハンディープロセッサを使います。

代替可 (別の食品、手法などでもよいこと)

- ・ もし半個ぐらいでしたら、手で搾って頂いても結構です。

留意事項 (料理のコツ、注意すべき事項など)

- ・ こうしますと味の染みが良くなります。

その他

- ・ なんか食べ物の絵ってこう、心が和んでとてもいいですね。

2.2 作業の階層構造

料理を作るための一連の作業は階層的な木構造としてとらえることができる。例えば図1のように、「豚肉の唐揚げの香味ソース添え」という料理は、肉の味付け、タレ作り、揚げる、などの作業に分割でき、タレ作りの作業はネギ、ショウガなどを切り、調味料を入れるなどの作業から成り立つ、という構造である。

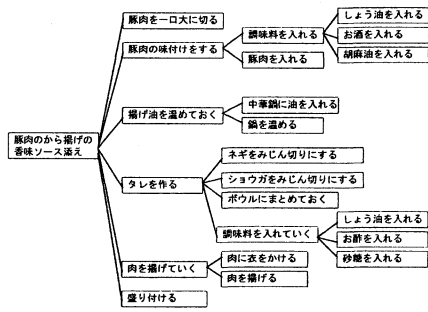


図 1: 料理作業の階層構造

このような木構造において、葉ノード (個別作業) 以外のノードに対応する発話は、「では、ステーキの材料にかかります」「まず、キャベツです」のようにそのノード以下の部分木の作業を始める宣言となる (本研究では作業宣言とよんでいる)。

一方、葉ノードとその親ノードからなる一段階の部分木は、あるひとまとまりの個別作業列を表すが、その個別作業の関係は次のような3つのパターンに分類することができる。ただし、最初の括弧の表現は、親ノードに相当する作業宣言である。

別材料同一動作

(調味料を入れていきます) 砂糖を入れて、醤油を入れて、ミソを入れます。

同一材料一連作業

(キュウリです) キュウリの皮をむいて、縦に切つて、刻みます。

準備必要型作業

(炒めていきます) 油をひいて、材料を入れて、炒めます。

2.3 複文の扱い

ここまでの議論では、各発話 (書きおこし文の句点ごと) を単位と考えてきた。ここでは、一つの発話が従属節を含む複文である場合について考える。

実際の料理教示発話データを調べたところ、多くの従属節は先に述べたように個別作業の程度、理由などを説明するものであり、それらは主節とまとめて一単位として取り扱って問題ない。

これに対して、発話を分割して扱うべきものとしては、基本的には「～けど (も) / ～けれど (も) / ～が」

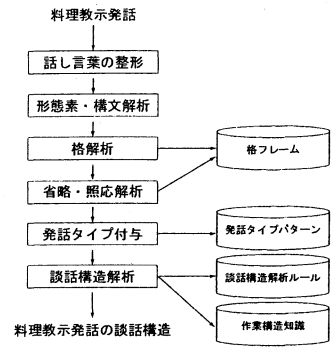


図 2: 料理教示発話の処理過程

の従属節だけであることがわかった。このタイプの従属節は別単位として扱い、作業の階層構造の解析で扱うこととする。

3 料理教示発話の基本解析

本研究における料理教示発話の処理過程を図2に示す。本節において省略・照応解析までの基本的言語解析について説明し、次節において談話構造解析について述べる。

3.1 料理教示発話の整形

まず、料理教示発話の形態素・構文解析を行う。我々の形態素解析システム (JUMAN)、構文解析システム (KNP) は、書き言葉テキストを対象として開発・改良してきたものである。しかし、たとえば新聞の引用文中の話し言葉などに対応するために口語表現にもある程度対応しており、間投詞の削除等の若干の前処理を行うだけで、そのままのシステムで解析を行うことが可能であった。

3.2 料理ドメインの格フレーム学習と省略・照応解析

日本語の文章では格要素が指示詞となったり、省略されることが頻繁に起こる。本研究で対象としているテキストはもとは音声で発話されたものであるため、よりその傾向が強いといえる。そこで、料理に関するテキストを WWW から収集し、料理ドメインの格フレーム辞書を自動構築した [3]。そして、構築した格フ

```

<WorkStructure id="3" label="炒める 2" type="準備
必要型作業" common="炒める">
  <Work id="1">フライパンを温める</Work>
  <Work id="2">油をひく</Work>
  <Work id="3">材料を入れる</Work>
  <Work id="4">炒める</Work>
  <Work id="5">味付けをする</Work>
</WorkStructure>

<WorkStructure id="4" label="切る 4" type="同一材
料一連作業" common="野菜">
  <Work id="1" type="opt">皮を剥く</Work>
  <Work id="2">縦に切る</Work>
  <OR>
    <Work id="3">細かく切る</Work>
    <Work id="3">菱形に切る</Work>
    <Work id="3">短冊型に切る</Work>
  </OR>
</WorkStructure>

```

図 3: 料理ドメインの作業構造知識

レームを利用して料理教示発話の省略・照応解析を行った(図4の例では、文中に括弧で示したものが、省略が補間された要素である)。

4 料理教示発話の構造化

4.1 発話のタイプ付け

2.1節で述べたように、料理教示発話は作業宣言、個別作業、料理状態、食品・道具提示、代替可、留意事項、その他に分類できる。

このうち、作業宣言、食品・道具提示、代替可、留意事項、その他については発話の文末表現のパターンを記述することで認識できる。具体的には、形態素列のパターンを記述し、各形態素について語、品詞、活用形、シソーラスの意味素性をチェックできる能力をもつものである。例えば、代替可については「～しても結構です／構いません／よい」など、留意事項については「～できます」「～しやすいです」「～を目安にしてください」などのパターンである。

個別作業、料理状態については、料理ドメインに対してすべての述語を列挙するというアプローチも考えられるが、他のドメインへの適用の可能性を考え、自動詞、形容詞+「なる」などを料理状態、それ以外を個別作業とする一般的な規則を用いている。

4.2 発話の談話構造解析

次に、前節の処理によって与えられた発話タイプを基にして、発話の談話構造を求める。モデルとしては、各発話を一つのノードとし、関係する発話がリンクされたグラフ構造を考える。料理教示発話のようなタスク指向の発話の場合、談話構造は、そのタスク、すなわち図1に示したような料理作業の木構造を反映したものが骨格となり、それを、代替可、留意事項、その他の発話が修飾するという構造となる。

しかし、本研究では図1のような多段の構造を求めることは対象外とし、葉とその上の一段階の階層、すなわち局所的な作業のまとまりを認識することを目標とする。すなわち、作業宣言と個別作業の並びの中からまとまりの検出を行うが、それを料理の作業構造に関する知識なしで行うことは難しい。そこで、2.2節で述べた、別材料同一動作、同一材料一連作業、準備必要型作業の3つの局所的な作業のまとまりに対して、具体的な作業構造知識を手手で記述した。

図3に作業構造知識の記述例を示す。表現はXMLの形式とし、<Work>...</Work>が構成要素の個別作業を示している。表現能力は正規表現と同程度の以下のものとした。

- type="opt" 0回または1回(オプションの作業)
- type="repeat" 1回以上の繰返し
- <OR>...</OR> その中のいずれかの作業を行う。1回以上の繰返しと組み合わせると、順不同のループ的作業が表現できる。

談話構造解析は、このような作業構造知識と、[4]の提案に基づく一般的な談話構造解析ルールを組み合わせで行う。図4に談話構造解析の例を示す。

5 実験と考察

5.1 データ

本研究ではNHKの「きょうの料理」50番組を対象として実験・検討を行った。キャプションデータには番組中の時間スタンプが与えられており、1番組は約200発話からなり、1発話の平均文字数は約20文字である。また、作業構造知識については、約50個の具体的な知識を作成した。

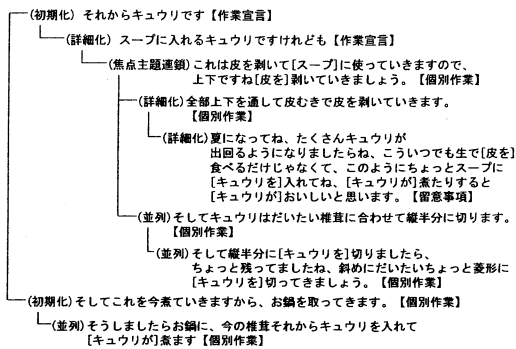


図 4: 料発教示話料の構造解析結果

5.2 解析精度と考察

料理番組 3 編について、番組中盤からそれぞれ 60 文ずつについて省略・照応解析，発話タイプ付与の精度を評価した結果を表 1 に示す。なお，省略・照応解析については解析結果が正しいものであるかどうか，すなわち適合率の評価とした。

省略・照応解析については，特にガ格の精度が悪い。他動詞のガ格はほとんどの場合省略されており，話者(料理作業者)が実際のガ格であるが，現在のパラメータでは文脈中から解を求めることが優先されており，多くの解析誤りを生んでいる。

発話タイプについては，適切な粒度の分類であり，表層的パターンで多くの場合正しく扱われている。ここで作業宣言・個別作業とそれ以外を正確に区別することが，次の談話構造解析で重要となる(談話構造解析は作業を中心に行われるので，発話タイプ付与はある意味で例外処理の役割をはたしている)。

談話構造解析については，作業構造知識を試験的に与え，テストを行っている段階であり，まだ定量的評価は行っていない。問題となるのは，作業構造知識に記述した作業と実際の発話での作業の表現のずれなどの問題であるが，基本的な枠組自体はうまく機能している。また，作業構造知識を人手で与えているということについては，将来的には自動学習する必要があると考えているが，当面は本論文のように，そのような知識が与えられたとしたら十分に使いこなすことができるかどうかについて研究を進める予定である。

表 1: 省略・照応解析と発話タイプ付与の評価結果

省略・照応解析		発話タイプ付与	
ガ格	25/68(37%)	作業	101/106(95%)
ヲ格	40/62(65%)	料理状態	1/2(50%)
ニ格	43/67(65%)	材料提示	15/17(88%)
計	108/197(55%)	代替可	5/5(100%)
		留意事項	19/21(90%)
		その他	47/65(72%)
		計	188/216(87%)

6 おわりに

さまざまな作業の対話的・映像的な自動教示システムを構築するための第一段階として，料理教示発話の構造解析を行った。話し言葉において問題となる省略・照応を自動学習した知識で解消し，作業に直接関係のない発話タイプを正しく認識し，さらに，トップダウンに与えたドメインの作業構造知識を用いて発話の談話構造を解析した。

格フレームを自動学習した点を除いて，特に新しい手法，アルゴリズムを適用したわけではないが，今後重要になると考えられる言語処理の応用分野において構造を正確に解析するというアプローチの可能性を示した。今後，実験規模を大きくして評価を正確に行うとともに，本研究で行った構造化を検索システムの観点から再度検討することが課題である。

参考文献

- [1] Youji Kiyota, Sadao Kurohashi, and Fuyuko Kido: Dialog Navigator: A Question Answering System based on Large Text Knowledge Base, In Proceedings of 19th COLING (COLING02), pp.460-466, 2002.
- [2] Hidekatsu IZUNO, Yuichi NAKAMURA, Yuichi OHTA. QUEVICO: A Framework for Video-based Interactive Media. In *Working Notes WS-5 International Workshop on Intelligent Media Technology for Communicative Reality, PRICAI-02 (Seventh Pacific Rim International Conference on Artificial Intelligence)* pp. 6-11, August, 2002.
- [3] 河原大輔, 黒橋禎夫: 用言と直前の格要素の組を単位とする格フレームの自動構築. 自然言語処理, Vol.9, No.1, 2002.
- [4] 黒橋禎夫, 長尾 眞: 表層表現中の情報に基づく文章構造の自動抽出, 自然言語処理, Vol.1, No.1, pp.3-20, 1994.