

命題間の接続関係に基づくマクロ構造の分析*

— マニュアルテキストを対象として —

レボウィッツ紀子(学籍番号 200721565)

研究指導教員:宇陀則彦

副研究指導教員:石井啓豊

1. はじめに

言語テキストがどのような談話構造を持っているのかを明らかにすることは、テキストの作成支援や自動要約、テキストからの知識情報抽出の質的向上につながる。談話というのはテキストひとつひとつの持つ内容の一貫性であるが、これはテキストジャンルごとに特徴的な構造があるとされている[1]。談話構造の分析手法には、文間の局所的な関係に着目するミクロ分析とテキストの構成要素(談話セグメント)がテキスト全体で担う機能を捉えるマクロ分析がある。ミクロ分析とマクロ分析を相互補完的に働かせることでテキスト全体の談話構造の把握が可能となる。このうちミクロ分析については、文間の接続関係や照応関係・修辞関係などを分析する手法が確立しているが、マクロ分析に関しては理論的な研究はあるものの、幅広く応用可能な分析手法が確立されているとはいえない。その理由として、テキストが多種多様であることに加え、それらの特徴把握には複数の視点があり、幅広く利用できる分析手法の考案が困難であることがあげられる。

そこで本研究では特定のテキストジャンルを対象として複数のテキストについて分析を行い、談話のマクロ構造抽出を試みる。

2. 談話構造のマクロ分析

談話のマクロ構造抽出への手がかりとして、まず談話構造のマクロ分析とはどのようなものを指すのかについて既存の研究をもとに具体化する[2]。テキストは談話セグメントと呼ばれる意味内容上のまとまりを持った文の集合から構成される。例えば、新聞等の解説記事であると、「主記」「解説」「背景」「意見」「見通し」の5つの談話セグメントから成り、主記を中心として背景と解説によって主題が構成され、意見と見通しがコメントとし

て主題に働きかけるという特徴的な構造がある。このように、テキストジャンルにおいて共通する談話セグメントを抽出することとそれらの関係を捉えることがマクロ分析には必要であるとされる。しかし、いかにして談話セグメントを抽出し、それらの関係を把握するのかといった点は一般化されていない。

そこで本研究では、表現上の構造がある程度定まっている家庭電化製品の取扱マニュアルを対象とし、見出しなどによりグルーピング化されている文の集合を利用して談話セグメントの抽出を行うこととする。その際に、マニュアルによって伝えられるべき製品の操作に関する知識に着目する。そのうえで、各談話セグメントの表す知識間の関係づけを行い、マクロ構造を得る。

3. 分析手法

本研究では、家庭電化製品のマニュアルテキスト4種全1860文を対象として、談話構造抽出のための分析を行う。はじめにミクロ分析である。テキストから2つの単文を抽出し、その文の核となる命題がどのような意味を持つのかを「ユーザの動作」と「状態の変化」という観点から分類する。次に単文間の接続関係を人手で抽出し、接続タイプを用いて分類する。接続タイプは市川[3]による接続関係の8分類にマニュアルにとって重要となる「順序」を加え、計9タイプを設定した。そのうえで、どのようなタイプの文がどのような接続タイプを持つのかといった接続のパターンを分析する。

次に、談話セグメントの抽出を行う。マニュアルテキストは内容を端的に表す見出しが存在するので、それを利用する。見出しごとのミクロ分析結果から出現頻度の高い接続パターンによって表されている事柄の共通性に着目し、談話セグメントを抽出する。そのうえで、抽出された談話セグメント間の関係を捉える。本研究では、特に製品使用の場面という視点から関係づけを行い、その結果得られた構造をマクロ構造とみなす。

* “Analysing Text Macrostructure based on Proposition Cohesion: the case of instruction manuals” by Noriko LEBOWITZ

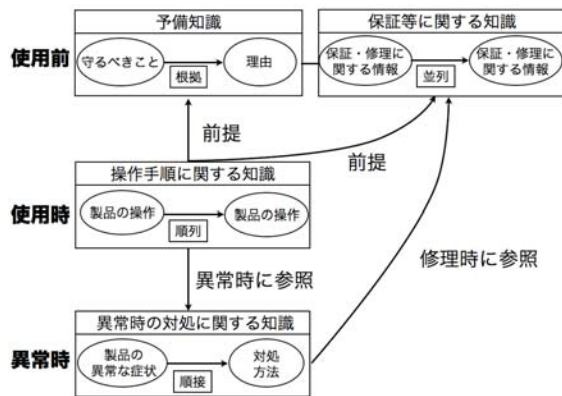


図1: マニュアルテキストのマクロ構造

4. 結果

4つのマニュアルテキストを分析した結果得られたマクロ構造を図1に示す。4つのマニュアルに共通する談話セグメントは「製品の操作に必要な予備知識」「製品の操作手順に関する知識」「製品異常時の対処に関する知識」「製品の保証やアフターサービスに関する知識」の4つであった。また各談話セグメントにおいて、出現頻度の高かった接続パターンによって述べられている主要な事柄は図1に示すとおりであった。例として、「製品の操作に必要な予備知識」において出現頻度の高かった接続パターンとその出現割合を表1に示す。

表1: 接続パターンの分析結果(予備知識)

前文のタイプ	後文のタイプ	接続	テキスト 1	テキスト 2	テキスト 3	テキスト 4
ユーザの動作: 誤 (要求+否定表現)	ユーザ・製品の状態: 異常 (演述)	補足	18.3	16.7	30.6	32.8
ユーザの動作: 正 (要求)	ユーザ・製品の状態: 異常 (演述)	補足	7.3	7.1	1.6	4.5
ユーザの動作: 誤 (演述+条件表現)	ユーザ・製品の状態: 異常 (演述)	順接	3.7	3.6	8.1	9.0

(%)

1つめから順に例文をあげると「製品の改造はしない→感電する」「トレーを必ずセットする→火災の原因になる」「他の機器と併用すると→火災の原因になる」となる。これらはいずれも前文でユーザの動作について何らかの命令を行い、後文でその根拠となる理由を提示しているものである。前文で命令している内容は製品を安全に使用するために守るべきことであり、後文ではその理由としてユーザ・製品の状態が異常になるためだと述べている。つまり、表現に基づく接続パターンは異なるが、これら3パターンによって表現されている事柄は共通しているということになる。

最後に、これら4つのセグメントをユーザによる製品操作の場面(使用前, 使用時, 異常時)という視点から

関係づけた。使用時に必要となる「製品の操作手順に関する知識」を中心とし、「製品の操作に必要な予備知識」と「製品の保証やアフターサービスに関する知識」を前提知識として位置づけた。それに加えて、製品の異常時に「製品異常時の対処に関する知識」をまた修理時に「製品の保証やアフターサービスに関する知識」を参照するという関係づけができ、最終的に図1に示すようなマクロ構造が得られた。

5. 考察

従来、マクロ構造とはテキストの要約やテキストジャンルごとに典型的な構成要素の把握であると考えられてきた。しかし、実際には得られる構造自体が問題なのではなく、テキスト分析の際に必要な視点を明らかにすることが重要であるとする。つまり、マクロ構造を得るための具体的な操作内容を定義することが必要である。

本研究では接続パターンの多様性を抽象化するという視点をもってマクロ構造の抽出を行った。しかし、接続パターンの多様性は、単に表現の多様性に基づくものではなく、文や接続関係を分類する際の視点の置きかたや1つの談話セグメント内に複数の要素が存在する際にセグメントをどう扱うのかといったことも含んでいる。これらの問題は、表現構造の特徴が見えにくいジャンルにおいてはより顕著となる。したがって、マイクロ分析結果を抽象化する際の具体的な操作内容については今後より詳細な検討が必要であるとする。

6. まとめ

本研究では、マニュアルテキストを対象として、その大局的な談話構造であるマクロ構造の抽出を行った。今後の課題は、談話セグメント抽出の精度向上とジャンルの枠を超えて幅広く利用可能なマクロ分析手法の提案である。

文献

- [1] 岸学編著. 文書表現技術ガイドブック. 東京, 共立出版, 2008, 152p., ISBN 978-4-320-00578-5.
- [2] 阿部純一ほか. 人間の言語処理—言語理解の認知科学—. 東京, サイエンス社, 1994, 343p., (Cognitive Science & Information Processing, 22), ISBN 4-7819-0739-3.
- [3] 市川孝. 国語教育のための文章論概説. 東京, 教育出版, 1978, 236p.