

14章:

文章のジャンル判別に寄与する指標の  
研究

-専門日本語教育への応用-

石田栄美

九州大学附属図書館研究開発室

# 専門日本語教育とは

- 日本語教育は、言語教育の一つの領域
- 通常、日本語を母語としない人を対象として行われる日本語の教育
- 専門日本語教育の現状
  - 大学・大学院における専門分野の学習・研究を日本で行おうとする学習者は、限られた時間の中で論述的な文章を理解し、作成する能力を身につけることが求められる。
  - 教育現場では教育の「効率性」が要求され、専門日本語教育に携わる者にとって大きな課題となっている。

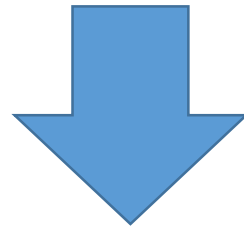
# 「話しことば」と「書きことば」

- 話しことば

- 去年X国で雨が全然降らなかった。天気がすごく悪かった。だから、みんなは食べ物がない。そして、びんぼうですごく困っている。だから、IMFはX国にたくさんのお金を貸してあげなくちゃいけない。

- 書きことば

- X国は昨年旱魃に見舞われた結果、食料不足に陥っている。したがって、IMFは貧困に苦しむX国に対して資金援助を行う必要がある。



- この違いはなにか？
- 判断するための多くの指標
  - 漢語語彙の選択、文の接続、適切な連語、慣用表現

# 研究目的

- 「書きことば」や「話しことば」などの形式を学習することは、母語話者だけでなく日本語学習者にとっても、文章のスタイルに対する意識を養うための有効な方法になりうる。
- 文章のスタイル＝「文体」＝「ジャンル」
- ジャンルとは「他の文章グループとの識別を可能にする何らかの共通の文体的特徴パターンをもった文章グループ」のこと
- 同一ジャンルに分類された文章には共通した特徴パターンがあるという仮説のもとに、具体的にどのような言語指標によってその文章ジャンルが分離されるのかということを実証すること。

# 研究方法

- 分析に用いるコーパス
  - 個人作成コーパス
  - 現代日本語書き言葉均衡コーパス (BCCWJ)
- 分析手法
  - 判別分析
  - 主成分分析
- 分析指標
  1. 複合辞 (接続語句と助詞相当句)
  2. 複合動詞の後項動詞
  3. 「手」のつく慣用句

# 用いたコーパス

1. 経済学教科書：総文数1124文
  2. 物理学論文：総文数2243文
  3. 工学論文：総文数1725文
  4. 文学論文：総文数4184文
  5. 経済学論文：総文数3889文
  6. 新聞社説：総文数6524文
  7. 文学作品
    - (1)近代文学作品：総文数2952文
    - (2a)現代文学作品：総文数1446文
    - (2b)現代文学作品：総文数2403文
- ・現代日本語書き言葉均衡コーパス (BCCWJ)

# 分析1：複合辞の指標とした場合

## • 複合辞

- 文の連結や文末表現等において形式化した語や助詞・助動詞が複合し、全体で一つの機能をもつ独自の表現形式のことである(例「ために」「はずだ」)。
- この複合辞は、日本語教育の現場では表現形式の有効な単位として用いられている。
- 接続語句と助詞相当句を選択
  - 接続助詞や接続詞等の接続語句が順接、逆接、添加というように文章の論理展開に重要な役割を果たし、論文に代表される論述的な文章において使用頻度が高いと推測されること
  - 助詞相当句についてもその一部が接続語句と重なっていて経験的にその使用傾向がジャンルによって異なる考えたため

# 分析1：指標としての接続語句と助詞相当句

- 市川(1978)、森田ほか(1989)を基準とし、選択
- 指標は、65語句

一方、～うえ[で]、うえ[に]、～から、～から～にかけて、しかしながら、したがって、すなわち、～ず[に]、～[た]結果、ただし、たとえば、他方、～ため[に]〈目的〉、～だけでなく、～つつ〈同時〉

等



# 分析1: 判別分析結果

- 分析対象7つの文章資料(7の2b以外)
- 各資料を一つのジャンルと仮定
- 各語句の出現率: 65語句の出現回数を調べ、一文あたりの出現頻度に換算した相対出現頻度
- 65語句の出現率を説明変数、ジャンルを基準変数として、ステップワイズ法を用いた判別分析を実施
- 19語句が選択された

①～において／におけるN、②したがって、③すなわち、④しかしながら、⑤～によって／によるN<方法>、⑥～とは<定義>、⑦～から～にかけて、⑧なお、⑨～について／についてのN、⑩～ながら、⑪～ので<理由>、⑫～から<理由>、⑬～ため[に]<目的>、⑭～によって／によるN<理由>、⑮よって、⑯～とともに<継起>、⑰ただし、⑱～として、⑲[それ]ゆえ[に]

图 14.1

## 表14.2

# 表14.4

# 分析1:主成分分析

- 対象コーパス
  - 4論文(物理学論文、工学論文、文学論文、経済学論文)計81編
- 語句の選択:17語句の選択
  - 各ジャンルごとにKruskal-Wallis(KW)検定で65語句の分布の違いを検討
  - KW検定結果で有意であった56語句から平均ランクが高い17語句を選択
  - 370編を対象に主成分分析を行った

①～において／におけるN、②したがって、③すなわち、④～によって／によるN<理由>、⑤なお、⑥～によって／によるN<方法>、⑦ただし、⑧～に関して／に関するN、⑨～として、⑩～ので<理由>、⑪～ため[に]<理由>、⑫～によって／によるN<根拠>、⑬たとえば、⑭～について／についてのN、⑮～のに対して<対比>、⑯～にもとづいて、⑰～に対して<対象>

# 分析1：主成分分析

- 370編の文章資料、17語句を用いて相関行列に対する主成分分析
- 成分数は固有値1以上の基準を設けた結果、5成分の初期解をもとめ、解釈のために、バリマックス回転を実施

# 表14.5

- 第一成分：提題・補足
  - 提題：「～に関して／に関するN」
  - 補足：「ただし」「なお」
- 第二成分：因果・方法
  - 「～ため[に]＜理由＞」「～によって／によるN＜理由＞」「～によって／によるN＜方法＞」
  - 「したがって」「～に対して＜対象＞」
- 第三成分：言い換え
  - 「たとえば」「すなわち」
  - 「したがって」「として」「～において／におけるN」
- 第四成分：根拠付け
  - 「～にもとづいて」「～によって／によるN＜根拠＞」
- 第五成分：不明
  - 「ので」
  - 「～のに対して」「したがって」「なお」



## 図14.2

第1成分「提題・補足」、第2成分「因果・方法」、第3成分「言い換え」、第4成分「根拠付け」、第5成分「なし」

# 分析2: 複合動詞の後項動詞を指標とした判別分析

- 対象コーパス: 7つの文章資料(⑦の2c以外)
- 文章のジャンル
  1. 論述文ジャンル(1~5)
  2. 新聞社説ジャンル(6)
  3. 文学作品ジャンル(7(1)(2b))
- 分析指標
  - 姫野(1999)を参考に複合動詞の後項動詞となる41動詞を選択
  - あう、あがる、あきる、あぐねる、あげる、あやまる、あわせる、いる、いれる、おえる、おくれる、おわる、あかる、かける、かねる、きる、こむ、こめる、すぎる、そこなう、そこねる、そびれる、そんじる、だす、たつ、たてる、つくす、つく、つける、つづける、でる、とおす、なおす、なおる、なれる、ぬく、のこす、はじめる、はてる、まくる、わすれる

# 分析2: 判別分析結果

- ステップワイズ法による判別分析
- 15後項動詞
- ①だす、②かける、③あがる、④かねる、⑤いる、⑥おわる、⑦なおす、⑧まくる、⑨いれる、⑩こむ、⑪きる、⑫つくす、⑬たつ、⑭つづける、⑮あがる

图 14.3

# 分析2: 判別分析結果

表14.6

表14.7

# 分析3:「手」を含む動詞慣用句と形容詞慣用句を指標とした判別分析

- 文章資料
  - BCCWJの書籍データ
  - 11分類
- ジャンル
  - 人文科学系(哲学、歴史、芸術・美術、言語、文学)
  - 社会科学系(社会科学、産業)
  - 自然科学系(自然科学、技術、工学)

# 分析3：指標としての「手」の慣用句の抽出

- 『基本慣用句五種対照表』(佐藤理史編, 2007)の慣用句3,628句
- 「手」が使われている動詞慣用句と形容詞慣用句を抽出
- 74項目

# 分析3: 判別分析結果

- 74指標を説明変数、ジャンルを基準変数としてステップワイズ法を用いて判別分析
- 三つのジャンルの判別に有効な、五つの慣用句
  1. 手に余る
  2. 手を打つ
  3. 手をこまねく／こまぬく
  4. 手をあげる
  5. 手を取り合う



## 表14.4

表14.9

表14.10

表14.11