

# 新聞投稿文の接続関係の分析

生田 和重

徳島文理大学 工学部

## 1 はじめに

学生の日本語文章力を高めるために、筆者が担当する授業において「分かりやすい実用文の書き方」を伝授している(生田 2002)。この学習では、最終課題として受講生全員が朝日新聞の声欄に E メールで投稿する。このようにして、過去 6 年間に蓄積した投稿文データは、学生の日本語文章力を把握し学習方法を改善していく上で、有効な素材であると考えている。そこで、文の長さや語彙の多様性などについて、統計的に分析したところ(生田ほか 2005a)、学生の作品と声欄に掲載された投稿文とは、有意な差があることが明らかになった。さらに、この結果をもとに、投稿前に確認すべき項目を設定し、それを実行させたところ、声欄への掲載率が大幅に改善された(生田ほか 2005b)。

このような学習方法の改善により、学生は文の簡明さや語彙の多様性を意識しつつ、文章を作成できるようになったと考える。しかし、各文間の関連性や文章全体の論理的な展開には、まだまだ改善の余地がある。そこで本研究では、学生の作品と声欄に掲載された投稿文を対象に、2 文間の接続関係を統計的に分析し、分かりやすく、説得力のある接続関係のパターンを見出すことを目標とする。

## 2 分析方法

### 2.1 分析対象データ

学生の投稿文は、E メールで朝日新聞編集局「声」係(dai-koe@asahi.com)へ送信される。そして、編集者の眼鏡にかなった投稿文は朝日新聞大阪版の朝刊に掲載される。そこで今回は、過去 6 年間に蓄積した学生の作品と朝日新聞大阪版の声欄に掲載された投稿文を分析の対象とした。なお分析の際

には、学生からの E メールを一括テキスト変換して得られたデータファイルと利用許諾契約を結んでいる「朝日新聞記事データ集 2002(朝日新聞社 2003)」を活用した。

### 2.2 分析の手順

文頭の接続詞や副詞などを手がかりに、文章の構造をコンピュータで解析する方法は多く提案されている(たとえば、竹内ほか 2001)。しかし、今回対象とした投稿文では、全文頭の 10 %程度にしか接続詞がついておらず、手がかり情報だけで全ての接続関係を把握することは難しい。また、手がかり情報を用いずに接続関係を判定することも検討されている(斉藤ほか 2006、山本ほか 2008)。しかし、この方法では大容量のデータを用いたルール学習が必要であるとともに、得られたルールを適用しても対象とした接続関係の 75 %程度しか把握できないようである。

そこで本研究では、文頭に接続詞や副詞などの手がかり語が存在する場合には、それらを参照して接続関係を把握し、それ以外の場合には手作業に頼った。当然ながら手作業では判定者によるバラツキが生じる。そこで、日本語文章力が高いと筆者が判断した日本文学科の学生 3 名にお願いして、合議の元に接続関係を判定してもらった。なお接続関係は、永野(1986)と阿部ほか(1994)を参考にして、表 1 のように設定した。

表1 使用した接続関係と関連する代表的な接続詞		
記号	接続関係	代表的な接続詞
A	展開型	だから、それで、すると、そこで
B	反対型	しかし、けれども、だが、ところが
C	累加型	そして、そのうえ、また、そうして、さらに
D	同格型	つまり、たとえば
E	補足型	なぜなら、というのは
F	対比型	または、あるいは、それとも
G	転換型	では、ところで、ときに、さて

	A	B	C	D	E	F	G	H
1								
2	二文間の接続関係							
3	No.	1→2	2→3	3→4	4→5	5→6	6→7	7→8
4	1	C	E	G	C	A	G	E
5	2	C	C	C	A	G	C	A
6	3	E	C	C	A	C	G	C
7	4	A	G	C	C	C	G	C
8	5	C	E	C	C	G	A	G
9	6	G	D	B	A	G	C	F
10	7	D	C	C	C	D	C	B
11	8	C	C	A	E	C	A	G
12	9	A	G	D	C	C	G	D

図1 接続関係データの一部

このようにして得られた接続関係データの一部を図1に示す。ここで、項目「No.」は投稿文の番号を示し、項目「1→2」は1文目と2文目の接続関係を表す。また、AからGまでの記号は表1に示すように接続関係と対応している。このようなデータ表を、学生の作品と声欄に掲載された投稿文について作成し、各投稿文内の接続関係の推移と接続関係パターンを定量的に分析した。

### 3 分析結果と考察

得られた結果をもとに両投稿文の接続関係の特徴を把握する。まず、図2と図3に学生の作品と声欄に掲載された投稿文の接続関係の推移を各々示した。ここで、横軸は接続関係を判定した2文の番号であり、たとえば「1→2」は1文目と2文目の間の接続を意味する。なお総文数には、当然ながらばらつきがあるので、最もデータ数の多かった総文数16文の場合を図示した。また縦軸は、接続関係(A,B,G)と接続関係(C,D,E,F)の比率を表している。市川(1988)を参考にすれば、(A,B,G)は「考えを掘り下げる接続関係」、(C,D,E,F)は「話題に広がりを持たせる接続関係」と考えることができる。

この結果より、学生の作品では「話題に広がりを持たせる接続関係」と「考えを掘り下げる接続関係」の出現頻度が単調に変化していることが分かる。すなわち、書き出しでは前者の出現頻度が後者のそれより大きく、その後後者の出現頻度が徐々に増加し、結末では大小関係が逆転する。一方、声欄に掲載された投稿文では、書き出しから結末に至る

までに、両接続関係の出現頻度の大小関係が周期的に変化している。このように両接続関係を織り交ぜることにより、文章に深みと説得力を持たせていると推察する。この結果を見ると、接続関係を判定してくれた学生が、「学生の作品は声欄に掲載された投稿文と比較して単調であり、接続関係を判定しやすかった」という印象を持ったことにも合点がい

く。筆者も実際に指導をしていて、過去形の文(過去の経験)が大半を占める単調な作文に出会うことがある。その際には、現在形の文(考えや事実)の例とそれを挿入する部分を提示するように心がけている。そして、修正前と修正後の文章を読み比べて、その違いを実感するように指導している。

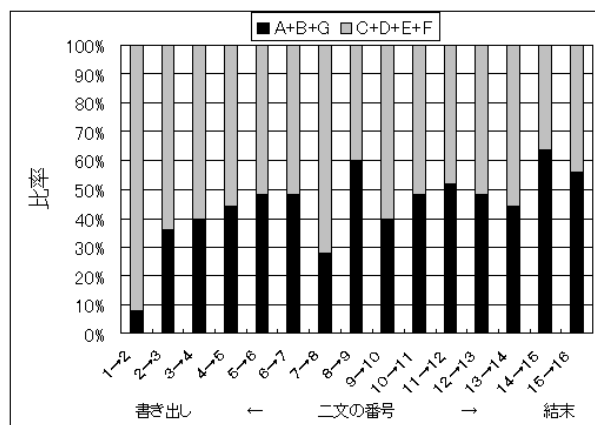


図2 接続関係の推移(学生の作品)

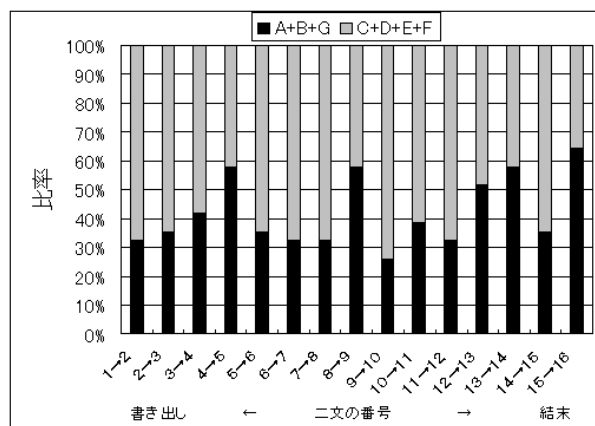


図3 接続関係の推移(声欄の投稿文)

係間の関連性（連続して出現する確率）である。図4はAに続く接続関係の出現率、図5はCに続く接続関係の出現率をまとめたレーダーチャートである。なお、図中の数字は出現率であり、以下の式で算出した。

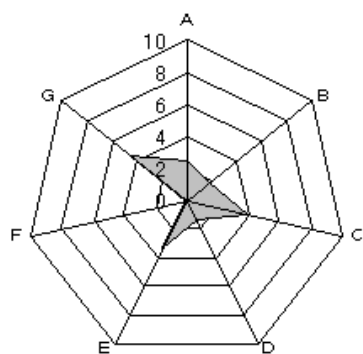
$$\text{出現率（\%）} = \frac{2 \text{ 接続関係が連続して出現する頻度}}{\text{接続関係の総数}} \times 100$$

まず図4では、A（展開型の接続関係）からC（累加型の接続関係）へと推移する場合に差が見受けられる。この場合には、声欄に掲載された投稿文における出現率は、学生作品における出現率の約1.5倍である。ここでAからCへの推移は、考えを掘り下げた後に関連する事柄を累加することを意味する。その頻度が多いことは、説得力や客観性を高めることにつながると推察できる。

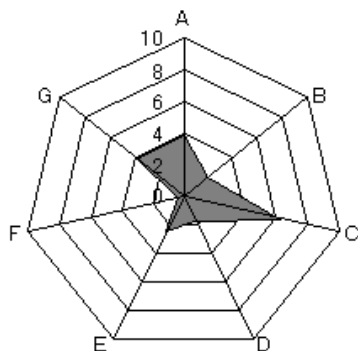
また図5では、C（累加型の接続関係）からA（展開型の接続関係）へと推移する場合に差が見受けら

れる。この場合には、声欄に掲載された投稿文における出現率は、学生作品における出現率の約2倍である。すなわち、前述の場合と逆に、関連する事柄を累加した後は、考えを掘り下げていると理解できる。さらに、声欄に掲載された投稿文ではCからCへの推移もCからAへの推移とほぼ同じ確率で出現している。関連する事柄を重ねて累加することにより読者の理解を深めることができると考える。ただし、累加型の接続関係が連続し過ぎると話が発散するので、声欄に掲載された投稿文のように、Cからの推移が多様であることは良い特性であろう。

今後の授業では、このような先輩学生作品と声欄に掲載された投稿文との差異を受講生に認識してもらい、接続関係を意識しつつ作文するように指導する必要がある。

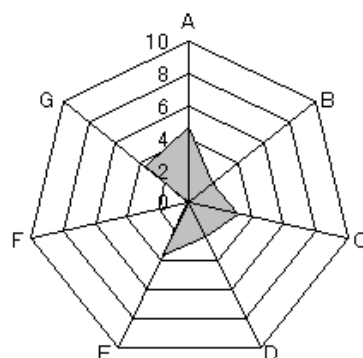


学生作品

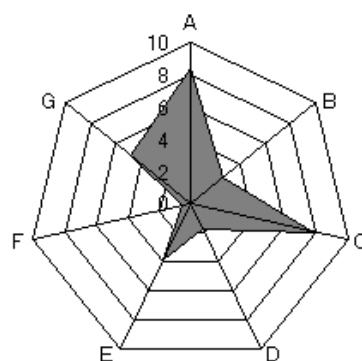


声欄に掲載された投稿文

図4 Aに続く接続関係の出現率  
(図中の数字：出現率（％）)



学生作品



声欄に掲載された投稿文

図5 Cに続く接続関係の出現率  
(図中の数字：出現率（％）)

#### 4 おわりに

接続関係を判定することの効果について、協力してくれた3名の学生にヒアリングした。その結果、「文章を読むスピードが上がった」、「論理の流れを意識しながら文章を書けるようになった」、「文章が硬くなったと先輩から叱られた（文芸部の学生）」などの評価をもらった。これは、1500件以上の投稿文を精読して接続関係を判定することにより得られた副産物であろう。授業では、これほど多くの文章を読んでもらう訳にはいかない。しかし、難解な接続関係を含む投稿文や説得力のある投稿文を選定して接続関係判定問題を作成し、「分かりやすい実用文の書き方」の授業で活用することは可能であると考え。今後は、学生の単調な文章を説得力のあるものへと変身させるように、具体的な学習内容を検討していきたい。

なお、本研究の一部は文部科学省科学研究補助金基盤研究（B）（課題番号 19300292、代表 石岡恒憲）の援助を受けた。

#### 参考文献

- 阿部純一ほか（1994） 人間の言語情報処理．サイエンス社，東京，pp.274-282
- 朝日新聞社（2003） 朝日新聞記事データ集 2002 学術・研究用．日外アソシエーツ株式会社，東京

- 市川保子（1988） 接続詞の用法と文脈展開 - 作文指導のための一試案 - ．筑波大学留学生教育センター日本語教育論集 ,3号 ,pp.175-185
- 生田和重（2002） 学内 LAN を活用した文科系学生に対する授業実施例．教育システム情報学会誌 , 19(1) , pp.28-32
- 生田和重ほか（2005a） 文科系学生が作成した投稿文の統計的な分析．日本教育工学会論文誌 , 29(1) , pp.35-42
- 生田和重ほか（2005b） 文科系学生が作成した投稿文の統計的な分析とその結果を活用した学習事例．日本行動計量学会第 33 回大会発表論文抄録集 , pp.382-385
- 永野賢（1986） 文章論総説．朝倉書店，東京，pp.104-108
- 斉藤真美ほか（2006） 大規模テキストを用いた 2 文間接続関係の同定．言語処理学会第 12 回年次大会 , pp.969-972
- 竹内和広ほか（2001） 新聞報道記事に対するテキスト構造解析の傾向とその手がかり情報の検討．計量国語学第二十二巻第八号 , pp.319-334
- 山本和英ほか（2008） 用例利用型による文間接続関係の同定．自然言語処理 ,Vol.15 ,No.10 , pp.21-51