

「決定木」分析によるコーパス研究の可能性: 副詞と共起する接続助詞「から」「ので」「のに」の 文中・文末表現を例に

玉岡 賀津雄†

SPSS 社の AnswerTree 3.0J に搭載された CHAID で決定木を描く統計手法が、コーパスから得られる 2 つ以上の変数からなる共起頻度を分析するのに有効であるかどうかを検討した。本研究では、3 種類の接続助詞「から」「ので」「のに」が、7 種類の副詞「何しろ」「何せ」「せっかく」「現に」「どうせ」「実際」「本当に」と共起する場合に、文中と文末の表現でどちらが使われるかを、新潮文庫 100 冊を使って解析した。決定木(図 1 を参照)は、3 種類の接続助詞と 7 種類の副詞の共起頻度によって、接続助詞の位置が文中にくるか文末にくるかを予測するという分析結果を描いたものである。決定木から 5 つの特徴を読み取ることができる。まず第 1 に、接続助詞「ので」と「から」が副詞との共起頻度において文中・文末で顕著な違いを示した。「ので」は文末ではほとんどみられず(5 回あるいは 4.59%)、「から」は頻繁にみられた(220 回あるいは 31.56%)。第 2 に、副詞「何しろ」と接続助詞「から」がもっとも典型的な文末表現であることが分かった(この種の組み合わせの合計 324 回のうち 140 回あるいは 43.21%)。第 3 に、接続助詞「から」と副詞の「せっかく」の共起は、かなりあるものの、文末では非常に少ないことも分かった(この種の組み合わせの合計 67 回のうち 6 回あるいは 8.96%)。第 4 に、接続助詞「から」と副詞「何せ」「現に」「どうせ」「実際」「本当に」はいずれも、文中・文末にはほぼ同じようなパターンで共起していることも示された。第 5 に、接続助詞「のに」は副詞の文中・文末の共起パターン(文中が 78.82%、文末が 21.18%)が、全体の共起頻度(文中が 72.73%、文末が 27.27%)と類似していた。以上のように、AnswerTree 3.0J による決定木の手法は、複数の変数からなる共起頻度データを構造的に分析することができ、今後のコーパス研究において有効な手段の一つとなるであろう。

キーワード: 決定木, SPSS AnswerTree, 共起頻度, 接続助詞, 副詞, 文末表現

Possibility of 'Decision Tree' Analysis on Collocation Frequencies: In the Case of Conjunctive Particles *kara*, *node* and *noni* Co-occurring with Adverbs at the Middle and End of Sentences

KATSUO TAMAOKA†

Using 'decision tree' drawn by a statistical algorithm CHAID in SPSS AnswerTree 3.0J, the present study investigated the collocation frequencies of three conjunctive particles (*kara*, *node* and *noni*) appearing in the middle or at the end of sentences, with seven selected adverbs (*nanishiro*, *nanise*, *sekkaku*, *gen'ni*, *doose*, *jissai*, and *hontooni*). Collocation frequencies were taken from the corpus of the *Shinchoo Bunko*

Collection of 100 Novels. Analysis results depicted in the decision tree predict two different particle positions (middle and end) of three conjunctive particles appearing with the seven adverbs. Five noteworthy collocation tendencies were observed. First, the conjunctive particles *node* and *kara* showed distinctive differences in the middle and ending positions when appearing with the adverbs: *node* was seldom seen at the end of sentences (5 times, or 4.59%), while *kara* was often seen at the end (220 times, or 31.56%). Second, the combination of the adverb *nanishiro* and the conjunctive particle *kara* occurred most frequently at the end of sentences (140 out of 324 times, or 43.21%). Third, although *kara* occasionally appeared with the adverb *sekkaku*, this combination was seldom observed at the end of sentences (6 out of 67 times, or 8.96%). Fourth, the conjunctive particle *kara* showed a similar pattern of collocation frequencies in the middle and at the end of sentences when combined with the five adverbs *nanise*, *gen'ni*, *doose*, *jissai* and *hontooni*. Fifth, the conjunctive particle *noni* appeared in the middle and at the end of sentences (78.82% in the middle and 21.18% at the end), similar to the overall percentages of the conjunctive particles (72.73% in the middle and 27.27% at the end). As such, in structurally depicting the collocation frequencies of conjunctive particles and adverbs, 'decision tree' analysis has considerable potential as a statistical approach in future collocation studies.

KeyWords: *decision tree, SPSS AnswerTree, collocation frequency, conjunctive particles, adverbs, ending expression*

1 はじめに

SPSS 社 (2001) は、AnswerTree 3.0J という統計解析用のソフトを開発した。この統計ソフトは、例えば、ある製品を購入する可能性のある顧客を見つける場合に、「購入する」と「購入しない」という 2 種類の名義尺度の値からなる目的変数 (あるいは、従属変数) を、顧客の多様な特徴である説明変数 (あるいは、独立変数) から予測する「決定木 (decision tree)」を描いて、製品を購入する可能性の高い顧客層を視覚的に特定してくれる。あるいはまた、銀行ローンを例にすれば、期限までに「返済する」と「返済しない」という値からなる目的変数を設定し、顧客の様々な特徴を説明変数として、ローンを返済する可能性の高い顧客 (あるいは、逆に可能性の低い顧客) を特定することができる。このように、AnswerTree は、これまでマーケティングなどの分野で使われることが多かった。一方、コーパス研究では、共起頻度のデータ分析について、名義尺度のデータを扱うため、ノンパラメトリックな統計であるカイ二乗を使った独立性の検定や母比率の検定の方法が使われることが多い。ところが、これらの検定法では、変数が 2 つ以上になると、有意な結果が得られたとしても、その原因が何に起因するかを特定するのが難しい。そこで、AnswerTree の統計ソフトを使って、複数の名義尺度の変数からなる共起頻度を分析して、決定木を描き、そのコーパス研究への応用の可能性を探った。

† 広島大学留学生センター、International Student Center, Hiroshima University

2 AnswerTreeのアルゴリズム

SPSS社が開発したAnswerTree 3.0J(SPSS 2001)には、4つのアルゴリズムが用意されている(詳細は、Breiman, Friedman, Olshen and Stone 1984; Loh and Shih 1997; SPSS 2001)。それらは、Chi-Squared Automatic Interaction Detection (CHAID)、Exhaustive CHAID、Classification and Regression Trees (C&RT)とQuick Unbiased Efficient Statistic Tree (QUEST)である。このうちのKass (1980)が開発したCHAIDとBiggs, de Ville and Suen (1991)が開発したCHAIDの修正版であるExhaustive CHAIDが、次の3つの理由で、コーパス研究への応用に適しているようである。なお、SPSS 13.0では、SPSS Classification Treesという統計解析ソフトで同じ分析をすることができる。

第1に、目的変数および説明変数に量的変数とカテゴリ変数(categorical data; 名義尺度と順序尺度の両方の変数を意味する)の両方を取ることができることである。コーパス研究では、カテゴリ尺度が使われることが多いので、このアルゴリズムは最適である。第2に、統計検定による有意水準を使用して、説明変数の値をすべて評価してくれることである。その際、目的変数に対して、統計的に等質であると判断された値を結合し、異質であると判断された値を保持することで、説明変数の値を最適に分類してくれる。目的変数が量的変数の場合には、F検定が使用され、カテゴリ変数の場合はカイ二乗検定が使用される。第3に、2進木を使用しないことである。AnswerTreeが決定木を描く場合に、樹木内の特定の階層が、C&RTおよびQUESTのアルゴリズムでは2分岐に限られるが、CHAIDとその修正版であるExhaustive CHAIDでは多分岐で描かれる。コーパス研究では、多様な表現の共起頻度を検討するため、多分岐であることが必須である。

3 本研究で使用するコーパス研究の事例

接続助詞は、通常は活用する語に下設して、「先行する表現を統括するとともに後行する表現に意味関係を示しながら続けていく。その点で、接続助詞は接続詞と機能が同じである」(秋元2001, p.388)。しかし、「接続詞が先行表現と後行表現とを分離したままで関係づけるのに対して、接続助詞は両者を関係づけることによって一続きにする点が異なる」(同)。確かに、「未成年だから、お酒は飲めません。」「雨が降るので、遠足は中止になった。」「高いのに、品質が悪い。」など、接続助詞は文中にくるのが普通である。厳密にこの定義に従うと、接続助詞が文末にくることはないという結論になる。しかし、「せっかく、いい天気なのに。」のように文末に接続助詞がくることもある。その際、たとえば「遠足は中止になった。」という事実など、この文が発せられるための文脈がすでに提示されている場合が多い。このような接続助詞で終わる表現は、特に口語においてよく見られることが報告されている(許2000; 白川1991, 1995, 1996)。その際、ある特定の副詞が、接続助詞の文末に共起する傾向があるとの指摘が滝沢(2002)によって

なされている。

そこで、接続助詞と副詞が共起する頻度を「新潮文庫 100 冊」のテキスト・ファイルを用いて調べた滝沢(2002)のデータを利用して、副詞と共起する接続助詞の文末表現の特徴を明らかにする。「何しろ」「何せ」「現に」「どうせ」「実際」「本当に」「せっかく」の7種類の副詞と共起する「から」「ので」「のに」の3種類の接続助詞について、これらが文中か文末でどのような頻度のパターンを示すかという3変数の条件で全体の出現頻度を一括して考察する。AnswerTreeによる解析では、接続助詞と副詞が「文末」または「文中」で共起するという2種類のデータからなる名義尺度を目的変数として、3種類の接続助詞と7種類の副詞という名義尺度からなる2つの説明変数で予測するという分析を試みた。

4 データの収集およびクロス集計

許(2000)が指摘している通り、接続助詞「から」などは、文末に生じる例がよく見られる。そこで、「から」が文末に生じている例を収集してみると、「何しろ」などの特定の副詞と共起している例が多々あることが分かった。これは文中用法の「から」と比較すると、著しく目立つ特徴のように思われる。また、接続助詞と共起する副詞として、「何しろ」以外にも「何せ」「現に」「どうせ」「実際」「本当に」などが頻繁に共起しているようである。そこで、共起例の多いこれらの副詞に関して滝沢(2002)のデータを使って調査することにした。

本研究のデータは、新潮文庫 100 冊から、「から」「ので」「のに」の3種類の接続助詞と「何しろ」「何せ」「せっかく」「現に」「どうせ」「実際」「本当に」の7種類の副詞が不連続で共起する頻度を検索したものである(詳細のデータおよび検索方法に関しては、滝沢 2002 を参照)。

もちろん、新潮文庫ではなく、毎日新聞 9 年分など他の資料を使うこともできる。しかし、新聞記事などは本稿で問題にするような「文末の接続助詞」のように「口語において頻繁に観察される現象」があまり多く含まれるとは言えない。そこで、口語的な表現も比較的多く含まれると思われる新潮文庫 100 冊を対象とした。新潮文庫には、安部公房の『砂の女』、遠藤周作の『沈黙』、井伏鱒二の『黒い雨』、井上ひさしの『ブンとファン』、石川達三の『青春の蹉跎』、松本清張の『点と線』、三島由紀夫の『金閣寺』、大岡昇平の『野火』、司馬遼太郎の『国盗り物語』、曾野綾子の『太郎物語』、立原正秋の『冬の旅』、高野悦子の『二十歳の原点』、筒井康隆の『エディプスの恋人』、吉行淳之介の『砂の上の植物群』など、さまざまな文学作品が収められている。したがって、新潮文庫に収められた 100 冊は、コーパスの規模として本研究の目的に照らして十分であると判断した。なお、形態素解析システム「茶釜」を利用して、新潮文庫 100 冊を分析したところ、全体の重なり形態素頻度は 109,716 であった。ただし、これは記号や表記の異なる場合(例えば、さくら、桜、サクラ)を含んだ合計である。また、全体の形態素の延べ頻度は、13,296,001 であった。

接続助詞と副詞の組み合わせは 21 通りである。さらにこれらの共起頻度は、接続助詞が文

中にくるか文末に来るかで分類した。例えば、「何しろポンコツだから、こんな急な坂は登れない。」であれば、「から」は、文中で出現するので文中用法である。一方、「何しろポンコツだから。」であれば、文末用法である。したがって、文中と文末の分類が加わって、組み合わせの可能性は全部で42通りとなる。共起頻度は、表1のクロス集計表にまとめたとおりである。ただし、終助詞の「ね」や「な」などを伴った「何しろポンコツだからね。」や、「何しろポンコツだからな。」という表現は、文末表現に含めて数えている。また、「ので」には、「んで」という撥音便があるが、これも含めて数えた。「んで」は、7つの副詞について、文中で4回、文末で1回だけみられた。撥音便「んで」の割合は、全体の4.59パーセントである。なお、前述のように、本研究では、日本語形態素解析ツール「茶釜」を品詞情報を得るために用いた。その際、「何しろいい天気なのだからね。」の「から」も接続助詞と同定されることに変わりはないので、「のだから」も「から」に含めて頻度を計算した。

表1 副詞と文中・文末にくる接続助詞の共起頻度のクロス集計表

接続助詞の位置	接続助詞	副詞						
		何しろ	何せ	せっかく	現に	どうせ	実際	本当に
文中	から	184	9	61	24	74	65	60
	ので	39	2	11	5	4	25	18
	のに	4	0	37	1	2	12	11
文末	から	140	5	6	10	21	20	18
	ので	3	0	0	1	0	1	0
	のに	0	0	10	0	1	2	5

表1をみる限りでは、「何しろ」と「から」の共起頻度が最も多く、文中で184回、文末で140回である。一方、「何しろ」が「のに」と共起して文末にくるような表現は見られない。また、「どうせ」と「から」の共起が文中で74回みられた。同用法が文末でも21回使用されており、文末用法としては「何しろ...から」に次いで2番目に多い。共起頻度が0である組み合わせは、文中用法では「何せ」と「のに」のみであるが、文末では、7種類の組み合わせにおいてみられた。共起しない組み合わせの数からみても、接続助詞は文末よりも文中で多く現れる表現であることが分かる。

以上のように、表1から大まかなパターンを読み取ることができるものの、3変数、つまり3次元での共起頻度を一括して適切に把握するのは極めて難しい。そこで、42通りの組み合わせ(7×2×3=42)を一括してより厳密に解析するために、AnswerTreeのCHAIDの解析による「決定木」の手法を使うことにした。

5 AnswerTree による分析

3 種類の接続助詞と 7 種類の副詞の共起頻度を説明変数として、文中か文末かという 2 種類の名義尺度の値を持つ目的変数を予測する分析を行った。

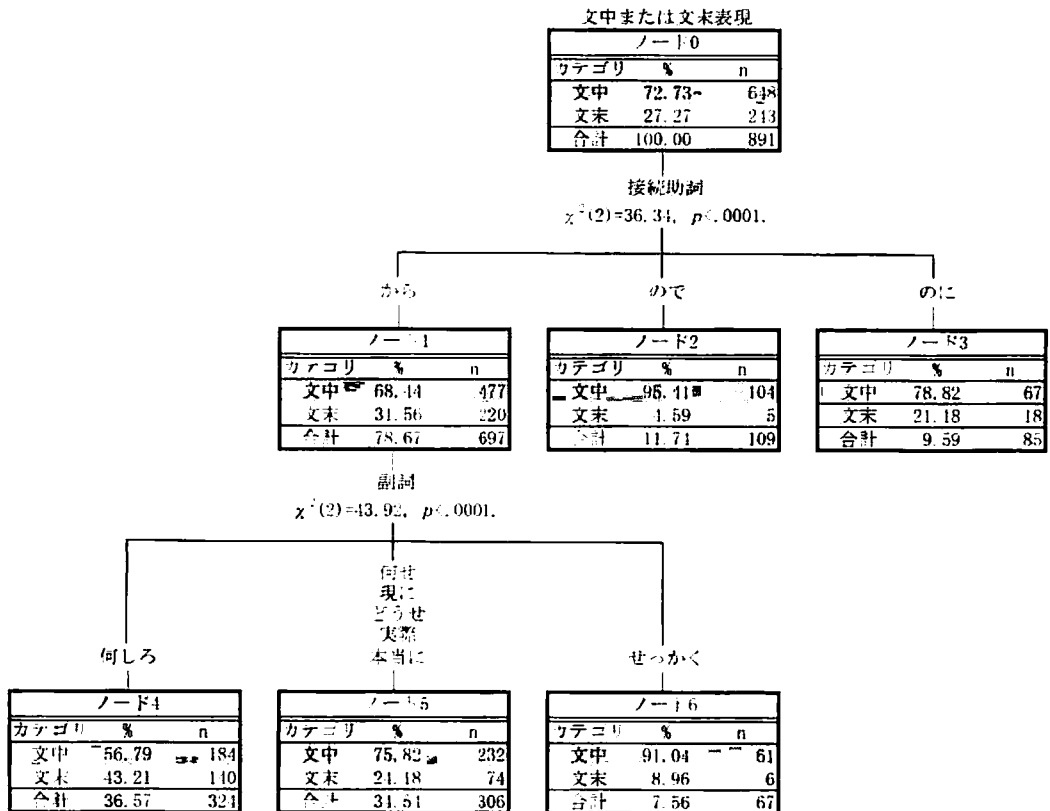


図 1 文中・文末で接続助詞と副詞が共起する頻度を描いた決定木

AnswerTree は、分析の結果を図 1 に示したような決定木の樹形図に描いてくれる。図 1 の決定木の出発点であるノード 0 は、対象となる接続助詞と副詞の組み合わせのうち文中で共起するものが 648 回であり、これが全体の 72.73 パーセントであること、一方、文末で共起するものが 243 回で、全体の 27.27 パーセントであることを示している。この文中・文末のノード 0 から伸びているのが、ノード 1 からノード 3 である。このことは、3 つの接続助詞がそれぞれ有意に異なった文中・文末の共起頻度を示していることを意味している [$\chi^2(2) = 36.34, p < .0001$]。さらに、接続助詞の「から」以外はそれ以上に樹木が成長していない。これは、接続助詞の「ので」と「のに」が、副詞との共起において文中・文末用法に特別な特徴を見出せないことを示

している。

表1の副詞との共起頻度をみると、(1)「ので」については、文中で104回も出現しているにもかかわらず、文末ではわずかに5回である。副詞と共起して「ので」が文末に使用されるのは稀であることが分かる。(2)「のに」については、副詞と共起して文末で使用されるのは18回あり、表1から分かるように、「せっかく」と共起するのが10回で、「本当に」と共起するのが5回である。しかし、他の副詞との関係は「どうせ」の1例を除いて共起ゼロで、7つの副詞との共起パターンを見出すことはなく、樹木が成長していない。一方、(3)「から」については、さらに樹木を成長させており、7つの副詞が3つに分類されている [$\chi^2(2) = 43.92, p < .0001$]。つまり、「から」については、共起頻度のパターンが副詞の種類によって異なることを示している。接続助詞の「から」は、副詞の「何しろ」と共起することが多く、合計でも324回であり、これは全体の共起頻度の36.57パーセントになる。「から」と他の副詞との共起頻度と比べて、この内訳が文末表現の特徴を顕著に示している。つまり、「何しろ…から」で終わるような文末表現は、140回もみられ、それは43.21パーセントになる。文中で両者が共起して用いられるのは、184回であり、これは56.79パーセントとなっている。このようなバランスは他の副詞との組み合わせではみられない。つまり、副詞と接続助詞とが文末で共起する典型的な形は、「何しろ台風がちかづいているから。」などによって例示される「何しろ…から」であることが分かる。

次に、表1では、一見すると「実際」と「から」が文末で20回も出現しているので、これが2番目だと考えがちである。しかし、文中での頻度を加味して分類すると、図1の決定木が示すように、「何せ」「現に」「どうせ」「実際」「本当に」の5つの副詞と「から」の共起頻度は、同じ分類であると判断されている。最後に、「せっかく」と「から」の共起は、文末ではわずかに6回(8.96パーセント)であり、文中では61回(91.04パーセント)と大きな違いを示している。そのため、「せっかく」は副詞のなかでも他と異なる分類となっている。

6 共起頻度からみた副詞と接続助詞の文末表現

接続助詞と副詞との共起頻度の傾向を文中・文末表現の基準で総括的に分析することが容易に行えなかった。本研究ではAnswerTreeに含まれているCHAIDのアルゴリズムによる統計的解析の手法を使って、接続助詞「から」「ので」「のに」が副詞と共起する場合に、文中と文末でどのような違いがあるかを電子化された新潮文庫100冊から全体的に解析した。AnswerTreeの統計ソフトは、図1のように、文中・文末でみられる接続助詞と副詞の共起頻度の最適な組み合わせを、視覚的な樹木のかたちの決定木で分かりやすく描いてくれるので、今後のコーパスに基づく研究に大いに活用できるのではないかとと思われる。本研究の結果は以下の5つに要約できよう。

第1に、接続助詞「ので」と「から」は、いずれも確定条件を表しているにもかかわらず、決定木では、副詞との共起頻度において文中・文末で顕著な違いを示した。「ので」と「から」

は、副詞との全共起頻度において、「から」が697回で「ので」が109回と違いがある。これを比率でみても両者の違いは顕著で、「から」が副詞と共起して文末でみられたのは31.56パーセント(220回)であるのに対し、「ので」はわずかに4.59パーセント(5回)に過ぎない。また、両接続助詞が現れる全頻度(滝沢 2002)でも、「ので」の文末表現は1.06パーセント(文末が74回で、文中が6,893回)であり、「から」の文末表現は17.81パーセント(文末が3,340回で、文中は15,377回)と大きく異なっている。決定木では、「ので」は副詞との共起についてノード2から樹木が伸びておらず、これは7種類の副詞との共起を文中・文末で議論するほどの頻度ではないことを示している。一方、「から」は、さらにノード1から樹木が伸びており、副詞との共起が文中・文末で頻繁にみられた。

これにはいくつかの理由が考えられよう。まず、「ので」の方が「から」よりも丁寧な表現であるため、倒置が起りにくく、その結果として文末に来ることがほとんどみられないのだと考えられる。例えば、「病院へ行きますので、明日仕事を休ませていただきます。」という表現と、これの倒置表現である「明日仕事を休ませていただきます。病院へ行きますので。」を比較した場合、後者の方が丁寧さという点において若干劣るという印象を受ける。これは、倒置という正順ではない形を用いているのと、結論を先に述べていることが原因であろう。したがって、丁寧であることが要求される場面では、「ので」が文末で用いられることは少ないのであろう。また、「熱があるので学校を休むから、伝えておいてください」のように、「から」は「ので」を含むことはできるが、その逆はできない(仁田 1982)ことも、大きな理由であると考えられる。

第2に、決定木の分析では、特に、文末用法としては、副詞の「何しろ」と接続助詞の「から」がもっとも典型的であることが、決定木のノード4から分かる。具体的には、接続助詞「から」は、副詞の「何しろ」と共に文中では184回、文末では140回もみられる。このように、文中・文末をあわせて全共起頻度の半数近くになるのは、この「何しろ」と「から」の組み合わせのみである。理由を表す「から」を使用する場合には、その理由を強調するために「何しろ」という副詞をつけることが多いのではないと思われる。なお、「何せ」は、「何しろ」と類似した表現ではあるが、類似していることは一方を使うと他方を使う必要がないという「相互に排他的(mutually exclusive)」な関係になってしまうので、頻度としては同じ傾向を示さなかったのではないと思われる。

第3に、副詞の「せっかく」が「から」と共起してみられることがかなりあるものの、その頻度が文末では低い。決定木は、この組み合わせをノード6として、他の副詞との共起と異なるとしている。実際、「せっかく」と「から」が共起して文末でみられたのは6回に過ぎず、文中の61回に比べて少ない。決定木は、「から」が副詞と共起してみられることが多いことを示しているものの、「せっかく」が「から」と組み合わせられるのは比較的少ないことを示唆している。

第4に、決定木は、接続助詞「から」と「何せ」「現に」「どうせ」「実際」「本当に」の5つの副詞が、文中・文末にほぼ同じようなパターンで共起していることを示している。滝沢(2002)

は、「本当に」の全資料中における絶対頻度が1,836回であるのに対して「何せ」が31回であり、もともとの副詞自体の頻度に大きな差があることを報告している。そうした違いは表1の共起頻度にも反映されるが、決定木はそれを加味して、文中・文末における「から」との共起においては、5つの副詞が同じようなパターンで出現していることを示した。

第5に、接続助詞「のに」は副詞の文中・文末の共起パターン（文中が78.82%、文末が21.18%）は、全体の共起頻度（文中が72.73%、文末が27.27%）と類似している。決定木のノード3から樹木が伸びていないことからわかるように、接続助詞「のに」は特定の副詞と共に用いられることが多いとまでは言えないという結果である。

7 AnswerTreeの描く決定木のコーパス研究への有効性

本研究では、共起頻度のデータをAnswerTreeのCHAIDのアルゴリズムを使って決定木を描いた。この手法が、コーパス研究に有効である理由は、本研究の例から、以下の3つに要約できよう。

第1に、本研究では接続助詞と副詞が共起して起こる文末表現について記述することが目的であった。図1の決定木をみると、この種の共起は、基本的に文中表現として出現していることが分かる。しかし、その中でも「何しろ...から」で終わる表現は、140回の頻度で、これは共起頻度全体の43.21パーセントとなり過半数に迫ることが決定木から分かる。つまり、文末表現の情報が、決定木で視覚的に分かりやすく描かれている。この平易で総括的な情報提示こそが、決定木が有効である理由である。

第2に、決定木が示す樹木の階層と分類である。3種類の接続助詞は、それぞれ異なった文末・文中での頻度を示している。「ので」と「のに」については樹木が下の階層に伸びていないが、これは、文末表現で使われる頻度が少なく、特定の副詞と共起して文末に出現していないことを示している。一方、接続助詞「から」は、副詞によって文末表現の共起頻度に違いがあり、「何しろ」は文末表現との共起が非常に多く、「せっかく」は極めて少なく、この二つは綺麗なコントラストを示している。その他の副詞は、その間に入り、すべて類似している。こうした一連の判断を自動的かつ最適に行ってくれるのは、決定木の大きな魅力である。

第3に、第1と第2の情報を、有意水準に基づいて提供してくれることである。本研究の場合であれば、目的変数が名義尺度であるためカイ二乗による統計的検定に基づいている。具体的には、図1のように0ノードから樹木が伸びて、7つの副詞が、最適なグループとして3種類に分類されている。しかも、変数の値を分類して決定木を描く根拠として、統計的な有意水準を使用している。これも、AnswerTreeのCHAIDの手法の利点である。

以上の3つの理由で、AnswerTreeのCHAIDの統計解析は、複数の変数で分類された共起頻度の分析にも有効であることが分かる。今後、コーパス研究で活用すれば、効力を発揮するものと思われる。

謝辞

本研究は、JST 社会技術研究「脳科学と教育」研究課題「言語の発達・脳の成長・言語教育に関する統合的研究」(研究代表者: 首都大学東京・萩原裕子)の助成により実施した。

参考文献

- 秋元守英 (2001). “接続助詞.” 山口明穂, 秋元守英 (編), 日本語文法大辞典, p. 388. 明治書院.
- Biggs, D., de Ville, B., and Suen, E. (1991). “A model of choosing multiway partitions for classification and decision tree.” *Journal of Applied Statistics*, 18, pp. 49-62.
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth.
- 許夏玲 (2000). 話し言葉の文末におけるモダリティ表現形式—「接続助詞」「条件形」「第二終止形」「引用助詞」. 名古屋大学大学院文学研究科日本語文化専攻博士学位論文.
- Kass, G. (1980). “An exploratory technique for investigating large quantities of categorical data.” *Applied Statistics*, 29, pp. 119-127.
- Loh, W. and Shih, Y. (1997). “Split selection methods for classification trees.” *Statistica Sinica*, 7, pp. 815-840.
- 仁田義雄 (1982). “助詞類各説.” 日本語教育学会 (編), 日本語教育事典 縮刷版, p. 399. 大修館書店.
- 白川博之 (1991). “【カラ】で言いさす文.” 広島大学教育学部紀要, 2 (39), pp. 249-255.
- 白川博之 (1995). “理由を表さない【から】.” 仁田義雄 (編), 複文の研究 (上), pp. 189-220. くろしお出版.
- 白川博之 (1996). “【ケド】で言い終わる文.” 広島大学日本語教育学科紀要, 6, pp. 9-17.
- SPSS (2001). *AnswerTree 3.0J user's guide*. SPSS.
- 滝沢直宏 (2002). “文末の接続助詞 (+ 終助詞) と不連続に共起し易い副詞—「から」の場合を中心に.” In Makino, S. (Ed.), *The Tenth Princeton Japanese Pedagogy Workshop: Proceedings*, pp. 132-142. Princeton, NJ: Princeton University.

統計ソフト

SPSS (2001). *AnswerTree 3.0J*. SPSS.

略歴

玉岡 賀津雄: 1990年、カナダ・サスカチュワン大学大学院博士課程修了(Ph.D.).
1990年から1998年まで松山大学講師および助教授。1998年から広島大学留

学生センター助教授を経て、現在、広島大学留学生センター教授。言語の認知処理モデルの構築をめざして、音韻、形態、語彙、統語構造がいかに関与しているかを解明するために幅広い実験研究を行っている。また、日本語のコーパスを使った統計的分析や中国語・韓国語・トルコ語・英語を母語とする日本語学習者の言語習得研究も手がけている。言語処理学会、日本言語学会、日本認知科学会、日本心理学会、日本教育心理学会、日本音韻論学会、社会言語科学会、日本語文法学会、日本音声言語医学会、日本語教育学会の各会員。

(2005年11月1日受付)

(2006年1月11日採録)