

# 体言止め表現に基づく感情推定と応答の自動生成

## Emotion Presumption and Automatic Response Generation Method based on Substantive Stop

鈴木 信雄†  
Nobuo Suzuki

津田 和彦†  
Kazuhiko Tsuda

### 1. はじめに

近年のインターネットの普及には目を見張るものがあり、ブログなどを使って個人が情報発信する機会が多くなっている。このような状況で使われる文章では、口語調で曖昧な表現が多く、思わぬ誤解などが発生し易い。そのような言語表現の一つに体言止めがある。体言止めは、修辞技法の一種で、名詞や名詞句のような「体言」で文章を終えることを言う。言い切らずに文の語尾に付ける終止形を省き、体言で止めることによって、余韻を残すことを指す。体言止めは、語尾の省略と共に、しばしば主語をも省略することがある。そのために、文の内容によっては曖昧となり、当該文のみを見ただけでは解釈が難しいものも多い。このような体言止めは、もともとは、俳句や短歌の技法だったが、近年では、各種の広告、新聞記事の見出しなどに多用されている。さらに、より感情を込めた強調表現や、携帯電話の狭小な画面サイズによる文字入力への煩雑さなどの理由により、最近のインターネット上での掲示板やブログなどに使われる口語調的な記述文章においても、体言止めが頻繁に出現している。

このような体言止めに着目した研究も、これまでにいくつか行われている。野畑らは、新聞記事に体言止めが多く現れることに着目し、記事の分類のためのイベントとして抽出している[1]。また、山本らは、新聞記事を対象として、体言止めを多用する新幹線車内の電光掲示板で使われている文章体への文章要約手法を提案している[2]。

本稿では、インターネット上に多く使われている口語的な記述文章における体言止め文に着目し、コーパスを構築して、その感情を理解する手法を提案する。さらに、本手法を用いることで可能となる詳細な感情分類をもとにした対話の自動応答手法についても提案する。

### 2. 体言止め用コーパスの構築

体言止めの感情推定ルール構築用に用いたコーパスは、携帯電話のインターネットからアクセス可能な複数のIT系企業における質問応答型のユーザーサポート用掲示板の文を収集して作成した。これらは、携帯電話のテンキーや小さい画面という制限されたユーザインタフェースのために、体言止めなどの曖昧表現が多く出現すると考えられる。文の数は全部で26,287文あり、この中で体言止めを含む文が1,206文であった。これらを使って感情推定ルールを構築した。

### 3. 体言止め用感情モデルの構築

体言止め文の感情を推定するためには、そのベースとな

る感情分類を定義する必要がある。本手法では、単なる情緒的な感情だけではなく行動や判断なども考慮しているPlutchikの感情分類[3]を拡張することにより対応した。今回構築したコーパスからPlutchikの分類で不足している「列挙」「助言」「説明」「呼びかけ」を追加し表1のような分類とした。

表1 体言止めの感情分類

感情分類	例
悲しい	はずれでしたか...残念。
列挙	▼自分が使用した有料サービスを確認する方法
助言	産婦人科はお金が高いから注意が必要
疑問	いつから暑中見舞い?
嬉しい	サイトも見放題♪
困惑	妹は『好きやから嫌や。でもどうしたらいいか分からん。』の一点張り。
拒絶	ていうか静電気ってやっぱりイマイチ。
謝罪	時間がないのでちょっと失礼。
説明	香辛料を組み合わせて、腸につめてポイル・燻製したもの。
呼びかけ	小村様。

### 4. 体言止めの感情推定ルール構築

各体言止め表現に対応する感情を推定することを考えると、体言止めの単語のみから、その意味を予測することは困難である。一般に、特定の言語表現には、それと共起関係にある表現を特定することによって出現するパターンを限定することが可能である。そこで、体言止めと共起関係にある形態素列を抽出することにより、感情推定に有効なルールを構築した。

まず、コーパスのデータに対して、体言止めを含む全ての文の形態素解析を行い、つづいて、体言止めの直前にある形態素列について、一意になると共に一定数以上の数量で共通して出現するものを抽出した。この体言止めと共起される形態素列を「体言止め句」と呼び、前項の感情分類に対応させて分類した。また、体言止め句を含む文の中で、文頭や文末に特定の記号により意図を一意に推定することができるものも見受けられた。そのため、これらを体言止めと関連付けてルール化した。さらに、体言止めの体言とは名詞のことであるが、記号やアルファベット文字列についても名詞と同じ役割を持つことから、体言止めの中に含めて解析を行った。

このようにして得られた体言止め句から感情を推定するためのルールの一部を表2に示す。本ルールで定義した体言止め句は57個、推定される感情は10個となった。

つづいて、これらの体言止め句の感情推定ルールを使って、インターネット上の体言止め文の感情を推定する手順を図1に示す。

† 筑波大学大学院 ビジネス科学研究科

表2 体言止めによる感情推定ルール(一部)

推定感情	体言止め句 ("":形態素の区切り)	例
列挙	助詞-格助詞-一般/動詞-自立 五段・ラ行 基本形/名詞	占いでわかること
助言	名詞-サ変接続/が/ひつよう	RSS リーダーにもファッションセンスが必要
説明	し/た/名詞-非自立-一般	腸につめてポイル・燻製したもの。
悲しみ	ざんねん	サブディスプレイがないので時間を見るとときに少し不便なのが残念
嬉しさ	すき	意外とアッサリで好き
困惑	いつてんばり	文句を言うと僕は知らないの一点張り。
拒絶	いまいち	サスペンションが堅くて、乗り心地いまいち。
謝罪	しつれい	横から、失礼。
...	...	...

Step 1	Web より 1 文を入力し、各単語の整形処理を行う。 具体的には、カタカナ表記の半角全角変換や文末にある記号の削除などを行う。
Step 2	形態素解析を実行する。
Step 3	文末の形態素が名詞であれば、以下を実行。
Step 4	文末より数えて 3 個の形態素を取得する。
Step 5	取得した 3 個の形態素を用いて、体言止め句を抽出する。
Step 6	抽出された体言止め句に対応するルールに従って、感情を推定する。

図1 体言止め文の意図推定処理手順

5. 体言止めに基づく自動応答

ここで、感情分類の「助言」に着目すると、さらに表3のように詳細な意図に分類することができる。これらの詳細な分類は、文章作成者の意図をより明確に示していることから、自動対話処理における相槌などの応答に利用することが可能である。表4に各詳細分類に対する応答を示す。これにより、ユーザサポートなどの自動応答業務への拡張が期待できる。

表3 体言止め句の詳細意図分類

体言止め句 ("":形態素の区切り)	詳細分類
ない/こと	注意喚起
かのう	教示
だいじょうぶ	激励
だめ	注意喚起
名詞-サ変接続/が/ひつよう	提言
も/名詞-形容動詞語幹	提言
名詞-サ変接続/が/(任意)	教示

表4 詳細意図分類による自動応答

詳細分類	自動応答文
注意喚起	わかりました。
教示	ありがとうございます。
激励	ありがとうございます。
提言	なるほど。

6. 評価実験

これまでに述べた体言止めを含む文の感情推定手法を評価するために、ルール構築用コーパスとは別に、携帯電話からも利用可能なインターネット上の質問回答サイトから6,755文のテキストデータを収集した。このデータの中で体言止め句を含む文は203文あり、ルールを適用可能な体言止め句を含む文は99文であった。この99文について本手法を適用したところ、適合率は100%となり、再現率は48.8%となった。この結果より、本手法の有効性を確認することができた。ここで、再現率が比較的低いが、これは適用できるルール数が少ないためであり、コーパスのデータ数を増やすと共に、ルールの数も増やすことで改善が可能である。表5に感情推定結果の例の一部を示す。

表5 評価実験結果(一部)

入力文	推定感情	判定
テレビ電話中の速度?	疑問	○
家族名義でも携帯電話のでもOK。	助言	○
ヘッダ情報が約5パケット	列挙	○
パケット割りスーパーって何	疑問	○
絶対に問い合わせはしないこと。	助言	○

さらに、今回作成したコーパスを使って、体言止めを含む文の使用傾向について確認した。まず、感情分類の上位を見ると、「列挙」が58.8%、「助言」が12.2%、「困惑」が1.5%となった。また、体言止めを含む文の文末の形態素種別では、「名詞-一般」が33.5%、「名詞-形容動詞語幹」が17.6%、「記号-アルファベット」が12.0%となった。これらより、体言止め文では、一般名詞が最も多く使われ、表現される感情は「列挙」と「助言」が多いことがわかる。

7. おわりに

本稿では、体言止めを有するインターネット上の曖昧文章の感情推定手法について述べた。本手法では、インターネット上の体言止め表現を多く収集したコーパスを構築し、Plutchikの感情分類を拡張することで体言止めに適した感情分類を提案した。また、体言止めと共起関係を持つ「体言止め句」と呼ぶ形態素列を定義して感情推定ルールを構築した。このルールを使って評価実験を行ったところ、高い性能を確認し、本手法の有効性を示すことができた。今後は、コーパスのデータを増やすことで再現率の改善を図ると共に、顔文字や三点リーダーなどの体言止め以外の曖昧表現を理解するインターネット上の総合コミュニケーションツールへ発展させる予定である。

参考文献

[1] 野畑周, 関根聡, 内元清貴, 井佐原均: “新聞記事における出来事を示す表現の分類と分析”, 言語処理学会 第10回年次大会 W1-4 (2004)

[2] 山本和英, 池田諭史, 大橋一輝: “「新幹線要約」のための文末の整形”, 自然言語処理 Vol.12 No.6 (2005)

[3] Plutchik R.: “Emotion: Theory, research, and experience, Theories of emotion”, Vol.1, N.Y. Academic (1980)