

情緒の生起する原因を表す日本語の用言の分布について

徳久雅人 田中努^{*1} 村上仁一 池原悟

鳥取大学 工学部 知能情報工学科

{tokuhisa, murakami, ikehara}@ike.tottori-u.ac.jp

あらまし 本稿では、情緒推定用の言語知識ベースの構築に向けて、情緒の生起する原因を表す日本語の用言の分布について調査した結果を報告する。情緒の生起する原因は、「情緒生起の特徴フレーム」に示される「特徴ラベル」を参照し、用言の語義は、「日本語語彙大系」に収録されている「結合価パターン」を参照する。結合価パターンに特徴ラベルを対応付けた結果によると、全結合価パターンの52.2%が表現の登場人物の情緒生起の原因を表すこと、そして、残りのパターンは表現に情緒主が登場しないまたは情緒生起の原因を表さないということが分かった。その結果から、基礎データとして、「特徴ラベル」や「用言意味属性」ごとに日本語の用言の分布を調査するとともに、用言に対応付けられる情緒の曖昧性について分析した。さらに、日記を対象とした情緒推定の実験を通じて、原因を表す用言を手がかりとする情緒推定の可能性について考察した。

キーワード 情緒, 感情, 情緒推定方式の分類, 語彙分析, 語彙知識ベース

1 はじめに

深い言語理解の一つとして、言語表現から話者や登場人物の情緒を推定することが挙げられる。情緒推定を実現する方法として、情緒推定用の語彙辞書を構築し利用する方法や、大規模コーパスにおける情緒的な表現との共起関係を用いる方法がある。語彙辞書の構築は、日本語の語彙に関する基礎的な知見を得るために有効な方法であり、また、最近では言語処理の高精度化に向けて語彙辞書・オントロジーについて見直しが始まり、語彙辞書への関心が高まっている。

情緒の語彙辞書についての先行研究として、形容詞・副詞・擬音語擬態語は飛田・浅田が網羅的にとりまとめ、情緒状態を明示する動詞は、一般のシソーラスにおいて「感情」のカテゴリに収録されている[1], [2], [3], [4]。助詞・助動詞およびそれらの相当語句や接続表現については、文型辞典として収録されていたり、表現を個別に深く議論されたりする[5], [6]。一方、言語表現に情緒状態が明示されていなくても、常識的な判断として情緒が読み取れることがあり、そのような深い理解も言語処理には要求される。たとえば、目良らは動作主や動作対象などの深層格と動詞の関係から好感度の変化する原因をとらえている[7]。遠藤らは日本語の依存関係に着目して Webドキュメントから情緒の生起原因を表す表現を大量に収集する方法を提案している[8]。土屋らは、連想関係を用いた常識推論の一つとして情緒推定を行っている[9]。しかし、これらの研究では日本語の語義について網羅的な検討を行っていない。

そこで、本稿では、情緒の生じる原因となる事態を表す日本語の用言の語義を網羅的に調査して、情緒の原因に関する日本語の用言の割合や用言に関わる情緒の曖昧性という基礎的な分布を調査する。

日本語の用言の語義や情緒の生じる原因は、それぞれの体系作りが大きな課題である。本稿では、日本語の用言の語義は、「日本語語彙大系」における「結合価パターン」で捉えることとする[10]。情緒の生じる原因は、徳久・岡田の「情緒生起の特徴フレーム」を利用する[11], [12]。「結合価パターン」は日本語の基本的な用言の語義を網羅している点に利点がある。「情緒生起の特徴フレーム」は原因の事態を的確な日本語のラベル(本稿では「特徴ラベル」と呼ぶことにする)で表しているため、日本語の用言の語義と比べることが比較的容易である。

そこで、本稿では、「結合価パターン」に「特徴ラベル」を対応付けることにより、情緒生起の原因を表す用言の語義を分析する。そして、その結果より、次の点を調査する。

- (a) 情緒生起の特徴フレームを基準とした分布
- (b) 用言意味属性を基準とした分布
- (c) 用言の表現から情緒を推定する上での曖昧性

まず、日本語の用言に対する基礎的な数値として、日本語の用言が情緒生起の原因をどの程度表しうるのか、た

*1 現在, (株)長府製作所 勤務.

たとえば、「喜び」と「怒り」とでは原因を表す用言の種類数にどれだけの違いがあるのかについて把握するために、(a)を調査する。次に、言語解析一般において、解析知識の汎用性を高める(スムージング)際に、用言の類似性を表すものとして用言意味属性が用いられることがあるので、本稿では、用言意味属性ごとに見て、情緒への関わりを把握するために、(b)を調査する。さらに、言語表現から情緒を推定する難しさは、表現と情緒の対応関係の曖昧性に依存するので、その把握のために、(c)を調査する。

また、特徴ラベルを対応付けた結合価パターンは、文とのマッチングにより情緒推定が行える。本稿は、情緒生起に関わる用言の分布を調査することが主たる目的であるため、考察として、手作業でパターンを用いた情緒推定を行い、情緒の原因に着目した情緒推定の可能性と問題点について確認する。

2 情緒推定の手がかり

本章では、言語表現から情緒を推定するというタスクについて整理し、本稿の位置付けを示す。

2.1 言語表現からの情緒推定の手がかり

日本語の表現は、表現しようとする対象を表す「客体的な表現」の部分、および、その対象に対する話者の態度などの「主体的な表現」の部分で構成されていると考えられているが、具体的な分類となると日本語の表現の分類には諸説ある[13]。本稿はその厳密な議論を行うものではないが、情緒推定タスクにおいて、話者の情緒を推定する上では主体的な表現が手がかりとなり、言語表現に登場する人物のもつ情緒を推定する上では客体的な表現が手がかりとなると仮定する。

別の視点として、文は話すためのものであるという考え方がある[14]。そこでの言語表現は、ある種の行為の性質があり、話し手や聞き手に推論を引き起こす、すなわち、言語表現の授受により対話者の心的状態が変化すると、考えられている。したがって、情緒推定タスクにおいて、言語の行為性が手がかりになると言える。

一方、言語表現を観測しても、その影響の及ばない立場の人がいる。たとえば、テレビドラマを見ている視聴者や小説の読者は、それらの架空の出来事(物語)を理解するが、それを真に受けて現実の自分自身だと思うことはない。しかしながら、視聴者・読者は喜びや悲しみを感じることもある。したがって、言語表現の物語性は、読者の情緒を推定する手がかりとなると言える。

以上を表1にまとめる。これらの4つのタイプは同時に言語表現に適用できる。たとえば、話者が1人称の代名詞により客体的な表現に登場する場合は#1と#2を同時に考慮することで、話者の情緒が推定できる。また、対話において対話相手の伝える内容に同情する場合は、#3と#4を同時に考慮することになる。

表1 情緒推定の手がかりの分類

タイプ	手がかり	情緒主	例文
#1	主体的な表現	話者	・雨に降られてしまった。(被害受け身) ・それはとても嬉しい。(感情表現文) ・とうとうあの土地が売れた。(修飾語句)
#2	客体的な表現	登場人物	・太郎が大学受験に合格した。(原因) ・太郎は喜んだ。(状態) ・太郎は万歳をした。(反応)
#3	言語の行為性	対話者	・手伝いましょうか?(提案) ・明日は晴れです。(伝達)
#4	言語表現の物語性	読者	・一杯のかけそばを3人で食べました。

これらの分類の元で関連研究の位置づけを示す。

タイプ1: 主体的な表現

主体的な表現とは、助詞や助動詞であるが、本稿では、話者の判断で形容する表現も主体的な表現とみな

す¹。第1章で述べた通り, [1], [2], [3], [5], [6]は, 日本語の形容詞, 形容動詞, 副詞, 擬音語, 擬態語, 文末表現, 助詞について議論している。文末表現や助詞については議論の余地が残されているが, 副詞・形容詞などについては情緒を表す語彙は網羅的に把握できており, 情緒推定のための知識ベース化が課題となる。

タイプ2:客体的な表現

情緒に関連する事態は次のように分けて考えることができる。

- (1) 情緒の生じる原因となる事態が発生する。
- (2) ある情緒の状態が生じる。
- (3) 情緒状態に対応する反応が生じる。

客体的な表現の表す事態が, (1)~(3)のいずれかに該当する際, 客体的表現の登場人物の情緒状態が推定できる。(1)に関しては, 第1章で述べたとおり, 目良らの好感度の計算[7], 遠藤らの用例ベースアプローチ[8], 土屋らの連想推論[9]が該当し, そのような表現を網羅的に把握することが課題として残されている。(2)に関しても第1章で述べたとおり, 一般的なシソーラスにおいて「感情」に分類される動詞が該当し, それは網羅的に把握できている。(3)に関しても, 同じくシソーラスにおいて「感情」のカテゴリに配置されており, たとえば, 「泣く」や「笑う」など「感情」に直結した反応を表す用言が見られる。ただし, 「(怒って,)机を蹴飛ばす」のように, 主たる語義が感情を表さない場合は, シソーラスには「感情」に分類されていないため, 今後の検討の余地がある。

タイプ3:言語の行為性

言語表現の授受により変化する対話者の状態は, 対話プランニングを通じて予測が可能である。Rosis は, 談話プラン知識における effect スロットに対話相手に生じる情緒を定義することにより, 医療説明対話における情緒の予測を行った[15]。しかし, 寓話や漫画を題材とした対話では対話行為や心的状態の認識が容易ではなく, 人手による心的状態のタグ付きコーパスの構築でさえ不安定である [16], [17]。

タイプ4:言語表現の物語性

言語表現に登場する人物への同情, 感情移入, 共感などを, 言語表現の物語性から推定する。このタイプは, 往住による分類がより詳しい[18]。ただし, 読者と登場人物の関係を+/-で表し, 掛け算的な変換により同情的か否かを求める方法が提案されているが[7], [19], 現在のところ言語処理技術として具体化できた点はない。

以上のように, タイプ1およびタイプ2の(2)については, 言語表現そのものは網羅的に挙げられており, 機械処理化に向け, 知識ベースを構築する時期にある。タイプ2の(1)は情緒推定の実現可能性が示されており, 日本語の表現における網羅的な調査を行う段階にある。タイプ2の(3), タイプ3, および, タイプ4では, 網羅的な言語知識を議論する段階には至っていない。

以上の状況において, 本稿では, タイプ2(1)の事態を表す日本語の用言について網羅的な調査を行う。

2.2 情緒生起の特徴フレーム

情緒生起の原因の分類について, Ortony らは, Goal, Preference, Standard を軸とした OCC モデルを示し, Elliott はそれを拡張して 26 種類の情緒を定義した[20], [21]。しかし, Goal, Preference, Standard という観点は抽象度が高いため, 用言との対応関係を分析するには思考上のコストが高く, 用言の分析漏れが懸念される。一方, 徳久・岡田は, 8 種類の情緒の生起する原因事態の特徴を約 120 種類に分類した[11], [12]。「獲得」や「賞賛」といった端的なラベルを用いてその体系を示しているのので, 用言と情緒生起の原因との対応関係を直感的に結びつけやすい。よって, 本稿では, 徳久・岡田の示した体系に従って, 用言に情緒生起の原因を対応付けることとする。

徳久・岡田は, Plutchik の pure emotion を参考に《喜び》《悲しみ》, 《好ましい》《嫌だ》, 《驚き》, 《期待》, 《恐れ》, 《怒り》の 8 種類を基本情緒とした。それぞれの生起するための原因となる事態の特徴を, 図 1 のような特徴フレームで定義した。この図は「喜び」の情緒の生起する原因事態の特徴をフレームで表現したものである。〈現状態

*1 形容詞・形容動詞は, 話者に内在する基準との比較により, 話者が単語を選択するため, 主體的と本稿では分類する。動詞の単語を選択する際も話者の視点によるものがあるが(たとえば, 「行く」と「来る」), 本稿ではその基準が話者の外からも判断できるので客体的とした。しかしながら, 本章では, これらの区別を厳密に追求することを狙ってはならず, 情緒推定の手がかりが整理できることを狙っている。

は前状態よりも好都合である)という特徴が最も抽象的な特徴(これを「最上位特徴」という)である。この特徴を持つ事象が原因となって《喜び》の情緒が生起することを示している。また、階層の深いところの特徴は、上位の特徴を継承したより具象的な特徴である。たとえば、《獲得》という最下位の特徴(これを「最下位特徴」という)は、《目標実現》の意味を継承しており、どんな物を獲得しても喜びが生起するというのではなく、「目標実現に必要なものを手に入れた」というときに《獲得》による《喜び》が生起することを示している。

なお、《心理的》や《目標実現》などは「中位特徴」という。本稿では、「最上位特徴」、「中位特徴」、「最下位特徴」をまとめて「生起特徴」と呼ぶことにする。また、《獲得》などの特徴名を「特徴ラベル」と呼ぶことにする。

後半での説明のために、最上位特徴の定義を表2に紹介する。

- (喜び/悲しみ: 現状態は前状態よりも好都合/不都合である
 生理的(内的な快/不快, 外的な快/不快)
 心理的(
 目標実現(
 情報収集(思惑どおり/思惑違い, 発見/見落とし, 判明/不明)
 計画(立案/無計画)
 実行結果(完遂/断念, 獲得/喪失, 有効/無効))
 対人関係(
 仲間意識(同意/反対, 同感/反感, 協力/非協力, 仲直り/仲違い)
 優劣関係(優越/劣等, 賞賛/非難, 服従/反抗, 厚遇/冷遇, 保護/見放す))
 その他)

図1 《喜び / 悲しみ》の特徴フレーム

表2 最上位特徴の定義

情緒名	最上位特徴の定義
喜び/悲しみ	現状態は前状態よりも好都合/不都合である
好ましい/嫌だ	有利/不利な状況になった
驚き	予測していなかった事象が起きた/予測していた事象が起きなかった
期待	好都合なことが起こると予測した
恐れ	危害を及ぼす事象・対象物を認識した
怒り	規則や当然のことが守られていない

3 日本語語彙大系における情緒生起の原因の分析

3.1 日本語語彙大系における結合価パターン

日本語語彙大系は、日本語の基本的な用言 6,118 個について、日英機械翻訳のためにその語義を網羅的に記述した結合価パターン 14,819 個を収録している。結合価パターンは、格要素と用言で構成する。格要素は、ある程度表現に自由度のある場合は一般名詞意味属性で制約された変数で記述し、慣用句のように表現が固定的な場合は字面で記述する。用言は字面のみである。

以下に「合う」という用言について例を示し、結合価パターンにより語義が区別されることを説明する。

#1 N1 (495 河川)が N2 (495 河川)と 合う

→ N1 join N2 (結合動作)

#2 N1 (*)が N2 (2449 条件 1176 需要 1361 願望・失望)とに 合う

→ N1 meet N2 (属性)

#3 N1(3 主体)が N2(3 主体)と 話が 合う

→ N1 and N2 have common topics (相対関係)

「合う」という用言には、34 種類の結合価パターンが存在する。ここには、そのうちの 3 つが示されている。N1 や N2 は名詞(句)とマッチする変数である。変数の後の数字とラベルは、変数にマッチしてもよい名詞の意味属性である。たとえば、「太郎が入社の条件に合う」という文とマッチするパターンは、#2 である。「太郎」は「河川」ではないの

でこの文は#1 にマッチしない。#2 は「が格」も「に格」もマッチするため、文全体がこのパターンにマッチする。#3 は「話が合う」という字面が必須であるためこの文はマッチしない。

日英機械翻訳を精度良く行うために語義の分解能が設定されているが、英語パターンの後ろに表示している「用言意味属性」からわかるとおり、日本語の語義を大きく逸脱するレベルの分解能ではない。ゆえに、結合価パターンの1つずつについて用言の語義をパターンから解釈することができる。言い換えると、結合価パターンの表す事態を想像することができるといえる。

3.2 結合価パターンへの生起特徴の対応付け

3.2.1 対応付ける内容

本稿では、以下の5種類の情報を「情緒属性」と呼び、そのセットを結合価パターンに対応付ける。ただし、「前提条件」は、主に簡単なラベルで記述し、変数との正確な対応関係は今後の課題とする。

- **情緒名:** 《喜び》, 《悲しみ》, 《好ましい》, 《嫌だ》, 《驚き》, 《期待》, 《恐れ》, 《怒り》
- **原因:** 情緒生起の特徴ラベル
- **情緒主:** 情緒を持つ人物。パターン中の変数で記述。
- **情緒対象:** 情緒を向ける対象。パターン中の変数で記述。
- **前提条件:** 補助的な条件

前節で示した#3 の結合価パターンに対して情緒属性のセットを付与した様子を以下に示す。

#3' N1(3 主体)が N2(3 主体)と 話が 合う

(a) 情緒名: 《喜び》 原因: 〈同意〉 情緒主: N1 情緒対象: N2 前提条件: N1 と N2 は仲間	(b) 情緒名: 《喜び》 原因: 〈同意〉 情緒主: N2 情緒対象: N1 前提条件: N1 と N2 は仲間
---	---

「話が合う」ということは、N1 が話したことを N2 が受け入れること、あるいは、N2 が話したことを N1 が受け入れることである。N1 の心境を考えると、N2 より〈同意〉が得られているので、それが原因となり《喜び》が生じると推定できる。ゆえに、(a)のセットがこの結合価パターンに対応付けられる。同様に、N2 の心境を考えると、(b)のセットもこの結合価パターンに対応付けられる。

3.2.2 対応付けの条件

先ほどの例より、結合価パターンへの情緒属性の対応付けについて様子が掴めるが、対応付けの作業を進める際に注意を払うべき点がある。それを以下にまとめる:

- 1) **登場人物の視点:** 人物を表すパターン中の変数に対して情緒主を仮定し、その人物の立場に立つこと。また、人物を表す変数が複数あるときは、各変数に対して立場を考えること。
- 2) **事態の解釈:** パターンの表す事態の解釈は次の点に注意すること。
 - 2-1) 時制, 相, 様相の指定がパターンに無い限り, 終止形で事態を解釈すること。
 - 2-2) 情緒主の視点で事態を解釈すること。
 - 2-3) 事態の解釈と情緒生起の特徴が一致するかどうかを判断すること。
- 3) **多義の付与:** 情緒名, 原因, 情緒主, 情緒対象が複数通り認められるならば, それぞれについてセットを作ること。1つのセットに絞り込むことはしない。

情緒主の視点で解釈することは、原因となる事態を見逃さないために、重要である。たとえば、「N1 が N2 を N3 から略奪する」という結合価パターンにおいて「略奪する」という言葉のイメージに頼ると「略奪された N3 の情緒《怒り》」だけに気をとられやすい。しかし、人物の視点を変数毎に置きながら事態を想像することにより、この用言では「略奪に成功した N1 の情緒《喜び》」に気がつく。

情緒生起の原因をあらかじめ定義しているため、結合価パターンが情緒生起の原因を表すかどうかの判断が、客観的に判断できる。ここで、パターンの表す事態が最下位特徴のラベルのいずれとも合わなくても、特徴フレームには最上位の特徴が定義されているため、最下位特徴に不足がある場合を補うことができる。その際、〈その他〉というラベルにより情緒名を対応付けることができる。また、結合価パターンの表す事態が抽象的であるならば、中位特徴のラベルを対応付けることもできる。

3.2.3 対応付けの工程

本稿では、前項で述べたとおり語義の解釈に客観性を持たせているが、作業漏れは避けられない。そこで、本稿

では2回の改訂を行う。具体的には次の3つの版を作成する。

第1版) 1人の作業者が結合価パターン辞書の先頭のパターンから順に情緒属性の対応付けを行う。

第2版) 再び同じ作業者が、前提条件の区別に注意しながら、見直しを行う。

第3版) 上記の作業者とは異なる者が検査して、対応付けに疑問のある箇所や追加案を作業者に示し、両者の協議により修正する。

3.3 対応付けの結果

対応付けの結果を以下にまとめる：

- **作業者：** 日本語語彙大系に収録されている全ての結合価パターンに対して、1名の作業者が情緒生起の特徴を対応付けた。第3版における検査者はテキストへの情緒タギングの経験者である。
- **日数：** 対応付け作業に要した時間は、1つの版に対して約3ヶ月であった。
- **対応付けの信頼性：** 対応付け作業に対する信頼性は、第2版を改訂して第3版を作るための検査の結果から、推察することができる。第2版から第3版への改訂において、検査者が指摘した箇所は、2,863カ所であり、検査者の指摘が妥当であるとして修正に応じた箇所が、2,494カ所であった。第2版では1パターンにつき1個以上の対応付けがあることから、17,587カ所が検査対象となっていた。対応付けに問題のあった割合が、14.2% (2,494 / 17,587)であった。第3版は指摘箇所が全て修正されているため、より信頼性が高いと推察される。
- **第3版の内訳：**
 - 日本語語彙大系には、6,118個の用言が収録されている。情緒属性のセットが対応付けられた用言は45.7% (2,798個)、対応付けられなかった用言は54.3% (3,320個)であった。
 - 日本語語彙大系には、14,819パターンが収録されている。情緒主が登場し、かつ、情緒生起の原因を表すパターンは52.2% (7,739パターン)であった。これらのパターンには情緒属性のセットが11,718セットが対応付けられていた。また、残りのパターン(7,080パターン)は、情緒生起の原因を表さないパターン、または、情緒主の登場しないパターン(たとえば、「M1が壊れる」)である。

第3版の詳細は、次章で説明する。

4 情緒生起の原因を表す用言の分布

4.1 情緒生起特徴を基準とした分布

4.1.1 情緒名ごとの分布

結合価パターンに対応付けた情緒属性セット(たとえば、第3.2節の(a)や(b))を情緒名ごとに集計した結果を表3にまとめる。この表より次のことが言える。

- もし、8種類の情緒に均等に対応付けがあったとすると12.5%ずつとなるのでそれを基準とすると、「好ましい」、「喜び」、「嫌だ」、「期待」は、その基準より多く、「悲しみ」はほぼ基準通り、「恐れ」、「怒り」、「驚き」は、基準より少ない。
- 「怒り」や「驚き」の数が著しく少ないのは、「規則や当然のこと」や「予測どおりでないこと/予測外のこと」という最上位特徴を1つの用言で表しにくいためであると考えられる。また、「恐れ」について、その最上位特徴は「危害」で構成されているが、それは「不都合」の下位概念であるため、対応付けられたセットの数は、「悲しみ」や「嫌だ」よりも少ないと考えられる。

表3 結合価パターンに対応付けられた情緒属性のセット数の情緒名ごとの分布

好ましい	喜び	嫌だ	期待	悲しみ	恐れ	怒り	驚き
2,641	2,296	1,907	1,820	1,356	1,106	507	85
22.5%	19.6%	16.3%	15.5%	11.6%	9.4%	4.3%	0.7%

次に、具体的に、情緒ごとに用言の例をみてみよう(表4)。用言の語義が格要素の語義に依存する場合は、格要素を補足して記載する。

- 「喜び」に関して、「惨敗する」は、「M1がM2に惨敗する」というパターンで、M2を情緒主とした場合である。逆に情緒主をM1とした場合は「悲しみ」が付与される。
- 「驚き」に関しては、もともと副詞が含まれているパターンや、偶然性を表す名詞が格要素に指定されている

パターンが目立つ。作業者の主観に依存して決められた部分と言える。たとえば、《驚き》の「N1(3 主体)がN2(2364 雨)に遭う」は、分析の作業者が「雨」に対して偶然性を感じているため、《驚き》が付与されたと思われる。同様に「喧嘩を買う」、「発掘する」、「目撃する」は事象に対する偶然性により判断されたと思われる。

- 《期待》に関しては、「他者に依頼する用言」、「他者と連絡をとる用言」、「ある事象の起こる前に行う動作」が目立つ。これらの用言の結果には新たな事態が予測されるため、《期待》の最上位特徴と一致する。
- 《恐れ》に関しては、身体的、社会的な危害の表される格要素が見られる。危害に対する体言の分類が、《恐れ》の原因事態の解析に有効であることが予想される。
- 《恐れ》の最下位特徴に〈叱責〉があるのだが、他者の《怒り》が予測される行為との関わりがあるので、《怒り》に関する用言と対比することで、より正確で漏れのない情緒属性の対応付けに繋がるとと思われる。

表 4 情緒ごとに見た対応する用言の例

情緒名	用言
《喜び》	愛護する, (条件に)合う, 軍配が上がる, (候補に)挙げる, 会得する, 援護する, 知恵を借りる, 手を借りる, 胸に応える, ご馳走する, 授かる, 惨敗する, 飼育する, 証明する, 助言する, 尽力する, 着服する, 満腹に成る, …
《悲しみ》	盗難に遭う, 喘ぐ, 頭が上がる(ない), 誤る, 争う, 焼きを入れる, 恨みを受け, 失う, 後ろ向きだ, 目を疑う, 訴える, 都を落ちる, 力を落とす, 思い切る, 夢にも思(ない), 折り合う, 抗議する, 降参する, ぶつかる, 見送る, …
《好ましい》	会う, 呼吸が合う, 水を開ける, (部屋/家屋に)上げる, (印象を)与える, 集まる, シャワーを浴びる, 安住する, 安心する, 気を利かせる, 口を聞く, 協和する, (人(職業)/労働で)食う, 駆使する, 決定を下す, 訓示する, …
《嫌だ》	喘ぐ, 呆れる, (授業で)当たる, 思案に余る, 怪しむ, 溝がある, 顔をうかがう, 支配を受ける, 深手を負う, (身体障害/病気類を)起こす, 気が置ける(ない), (敵・味方を)抑える, (支払い/料金を)溜める, 墮落する, …
《驚き》	(雨に)遭う, ばったり遭う, 手の内を明かす, (リストに)上がる, 不意を打つ, (目/口を)押さえる, 思いがけない, 喧嘩を買う, 裏を搔く, 不意を食う, 身につまされる, 発掘する, 舌を巻く, 見つかる, 目を見張る, 目撃する, …
《期待》	(資材を海外に)仰ぐ, (薬品を)仰ぐ, (立案を)明かす, 幕を開ける, 例に挙げる, のろしを上げる, (心当たりを)当たる, (関心/同情/話題が)集まる, 依頼する, 電話を入れる, 打ち明ける, 夢に描く, 演説する, お参りする, …
《恐れ》	毒杯を仰ぐ, 脅威を与える, (人間活動に)当たる, (強豪に)当たる, 違反する, 慰留する, 隠蔽する, 迂回する, 疎い, うろたえる, 追い込む, 覚悟する, 疑念を掛ける, 偽証する, (力/知恵/背丈を)比べる, (名称を)汚す, …
《怒り》	煽る, 欺く, 言い返す, 喧嘩を売る, 横領する, ぺてんに掛ける, 強いる, しくじる, 詐欺をする, 作り話をする, だます, 挑戦する, 不幸に付け込む, ミスに付け込む, 手出しする, 反論する, 仕事を回す, どじを踏む, 笑う, …

4.1.2 生起特徴ごとの分布

情緒生起の特徴は階層性がある。用言の語義が抽象的なことを表す場合には、その結合価パターンには、最下位特徴を対応付けるのではなく、中位や最上位の特徴を対応付け、また、最下位特徴に適切なラベルが無い場合は中位特徴や「その他」を対応付けるという方針をとった。そこで、その確認として、生起特徴の階層を見ながら、対応付けられた情緒属性セットの分布を調べる。

その結果を表 5 に示す。最上位特徴を対応付けた数は「H」の行に、中位特徴を対応付けた数は「M」の行に、最下位特徴を対応付けた数は「L」の行にそれぞれ記した。斜線部は、特徴フレームに該当する特徴ラベルが定義されていない部分である。内訳を以下に説明する：

- 11,718 セットがパターンに対応付けられ、最下位特徴の対応付けは 10,927 セットであった。その割合は、93.2% である。逆に、最上位特徴、中位特徴、その他の特徴の対応付けは 6.8% である。もし、最下位特徴以外の対応付けに不具合が無ければ、最下位特徴のラベルは適切であるといえる。
- 中位特徴〈生理的〉の下位の特徴には〈内的な快〉、〈外的な快〉がある。〈生理的〉に対応付けられた用言に

は、「全快する、治す、吹き返す、若返る」などがある。下位特徴の両方に共通する語義であるため、中位特徴に対応付けられた。

- 中位特徴〈情報収集〉は、《期待》においては、「探りを入れる、(事件/問題に)メスを入れる、インタビューする、(流行/跡/変動を)追う」という、『情報を収集する行為を開始する時点の事象』を表す表現であった。これらは『調査』に関連する用言であるが、《期待》の最下位特徴には該当する特徴ラベルが無かった。「期待の特徴フレーム」に追加する必要がある。
- 中位特徴〈実行結果〉に対応付けられた用言には、「配る(好ましい)、汚す(嫌だ)、ぶつける(悲しみ)」がある。「配る」は最下位特徴〈補助〉に該当する。「汚す」、「ぶつける」は最下位特徴〈喪失〉に近い。これらは改訂する必要がある。
- 中位特徴〈対人関係〉に対応付けられた用言には、「惚れる(期待)、熱愛する(期待)」という用言意味属性〔感情動作〕の用言が含まれていた。「惚れる」の直接表現は「好ましい」であるが、「惚れた後」には対人関係に関して様々な事象の発生が予測されることから《期待》と判断した。深く考えて対応付け作業が行われていたことが伺える。
- 〈その他〉に関して、564 セットが対応付けられた。幾つか特徴的なものを説明する。
 - 《好ましい》に関しては、「ドライブする、登山する、保管する、水上スキーをする、(たばこを)吸う」などがあり、個人の嗜好に依存している。これらは、逆に個人に依存して《嫌だ》と分類されても誤りではない。
 - 《嫌だ》に関しては、「(授業で)当たる、言い訳が立つ、遅刻する、年を取る、血が滲む、舌が回る(ない)」などであり、これらは平均的な人々が嫌だと感じる原因として妥当であると考えられる。
 - 《恐れ》に関しては、「隠れる、避ける、忍ばせる」などであり、隠れたり避けたりすることの生じた原因は《恐れ》であるため、これらの分類は直接的には「反応」の事象である。しかし、これらの事象は《恐れ》の解消にはなっておらず、「見つかること」や「避けられないこと」のように2次的な《恐れ》の原因になることがある。ゆえに誤りではない。

表 5 生起特徴ごとに見た情緒属性セットの分布

	喜び	悲しみ	好ましい	嫌だ	驚き	期待	恐れ	怒り
最上位 H	1	1	0	2	1	2	3	1
生理的 M	10	5	42	37	0	10	7	0
L	13	68	159	306	1	20	57	10
心理的 M	2	10	0	1	0	0	2	1
目標実現 M	0	0	0	0	0	0	0	1
目標 M			0	0				0
設定 L			454	561				95
情報 M	0	0	0	0	0	6	2	0
収集 L	158	57	294	67	59	276	283	13
計画 M	0	0				0	0	
L	113	134				432	219	
実行 M	0	2	3	1	0		0	0
結果 L	1149	488	642	250	13		4	126
対人関係 M	0	0	0	0		0	1	3
仲間 M	3	0	14	17		31	5	0
意識 L	212	122	548	349		471	299	156
優劣 M	0	0	0	0		0	0	0
関係 L	614	431	362	232		336	192	82
その他	21	38	123	84	11	236	32	19

さらに、最下位特徴の使用数についての順位を調べてみよう。表 6 には多く使用された最下位特徴を、表 7 には使用の少ない最下位特徴をそれぞれ示す。使用頻度の高いもので、「制限」と「自由」は目標設定という点で相反するが、ともに多く表れている。使用頻度の低いものは、真に日本語において該当する用言が少ないかもしれないが、

再度見直す必要がある。

表 6 用言に多く対応付けられた情緒生起の原因

順位	情緒・原因	セット数	用言の例
1位	嫌だ・制限	561	義務がある, 監督する, 禁じる, 風雪に耐える, …
2位	喜び・獲得	458	回収する, 確保する, (品質を)格上げする, …
3位	好ましい・自由	454	荷が下りる, 緩和する, 幅を効かせる, 独立する, …
4位	好ましい・補助	442	完備する, 矯正する, 陰影を付ける, 美化する, 養う, …
5位	期待・成算	432	(構想・計画を)暖める, 決心する, 決断する, 強気だ, …

表 7 用言への対応付けの少なかった情緒生起の原因

順位	情緒・原因	セット数	用言の例
1位	怒り・内的な不快	0	
	驚き・成功	0	
3位	喜び・外的な快	1	(寒暖を)凌ぐ
	驚き・五感急変	1	(目・口などを)押さえる
5位	怒り・脱退	3	裏切る, 独立する(2)
	怒り・責任逃れ	3	委任する, 迫る(2)
6位	嫌だ・一方的	3	重荷を負わせる, 牛耳る, 説き伏せる

以上より次のことが言える:

- 情緒生起の原因を, 用言で表すことについて
 - 《驚き》や《怒り》の原因を表す用言は少なく, 《好ましい》, 《喜び》, 《嫌だ》, 《期待》の原因を表す用言は多い.
 - 〈制限〉や〈自由〉と目標設定に関する原因を表す用言が多い.
- 個人の依存性について
 - 《好ましい》と《嫌だ》の例に見られた動作の嗜好性, 《驚き》の例に見られた事象の偶然性は, 個人への依存度が高い.
- 情緒生起の特徴フレームについて
 - 対応付けの全セット数のうち 93.2%で最下位特徴が使われていた.
 - 中位特徴の対応付けられた用言は, 適切な抽象度の用言であった.
 - 一部の最下位特徴については対応付けを再確認する必要はあったが, 全体的にみて, 特徴フレームは, 適切に情緒の原因を網羅しているといえる.

4.2 用言意味属性を基準とした分布

一般に言語解析用の知識を効率良く運用するために意味属性を用いて汎化したルールや用例を使用すること(スムージング)がある。意味属性を用いたスムージングの前提は、同一の意味属性の用言が同様の性質を持つという仮定が成り立つことである。ゆえに、情緒の推定において知識ベースのスムージングを成功させるには、同一の意味属性のパターンが同様に情緒の原因となりうるかどうかが問題となる。そこで、用言意味属性ごとに見て、結合価パターンが情緒に対応する割合を調査する。

表 8 に結果を示す。用言意味属性ごとに、対応したパターン数を、存在するパターン数に対する割合で記載する。たとえば、用言意味属性が〔2 状態〕である結合価パターンは 25 個であり、そのうちの 24%が情緒《好ましい》の生起する原因を表していることを示している。

〔13 心的状態〕は、《悲しみ》, 《嫌だ》, 《恐れ》などネガティブな情緒の割合が高いが、《喜び》が 22%とあり、一概にスムージングが行える状況ではない。全体的に見ても対応する割合が、10%台から 20%台である箇所が目立つ。これは、用言意味属性だけで情緒の判定が難しいことを表し、単純なスムージングには不適切であることを表す。用言意味属性の分解能が、情緒生起の特徴を区別するには足りないことを意味しているのだろうか。そこで、用言意味属性と情緒生起の特徴とが、そのラベル名から直ぐに対応付けられそうな「所有的移動」と「消滅・破壊」について詳しく見てみる。

表8 用言意味属性と情緒ごとに見た対応する結合価パターンの分布

用言意味属性	パターン数	喜び	悲しみ	好ましい	嫌だ	驚き	期待	恐れ	怒り	なし
1 事象	0									
2 状態	25	8%	4%	24%	8%					64%
3 抽象的關係	1			100%	100%					
4 存在	61	2%	2%	18%	7%					77%
5 属性	3,783	4%	4%	7%	8%	0%	2%	3%	1%	77%
6 所有	78	23%		45%	1%		9%	4%		37%
7 相對關係	363	31%	27%	21%	19%	0%	10%	8%	8%	23%
8 因果關係	48	6%	2%	10%	2%		4%	2%		83%
9 精神的關係	2			50%						50%
10 知覚状態	60		3%	22%	55%			2%		23%
11 感情状態	275	5%	25%	17%	20%	1%	17%	11%	3%	17%
12 思考状態	25	16%	8%	12%	16%		20%	16%	8%	24%
13 心的状態	9	22%	67%		11%	22%		33%		11%
14 身体状態	54		7%	17%	39%		4%	7%		31%
15 自然現象	155			1%	4%			3%	1%	95%
16 行動	159	18%	3%	16%	10%		30%	23%	6%	25%
17 物理的行動	1							100%		
18 物理的移動	511	8%	6%	19%	11%		12%	9%	1%	59%
19 所有的移動	409	63%	13%	32%	9%		19%	3%	6%	11%
20 属性変化	2,588	14%	9%	17%	12%	0%	6%	6%	3%	52%
21 身体変化	373	10%	13%	12%	30%		3%	6%	0%	36%
22 結果	269	29%	26%	9%	18%	6%	2%	15%	3%	
23 身体動作	1,811	12%	6%	17%	13%	0%	10%	8%	4%	50%
24 利用	121	18%	13%	20%	6%		17%	7%	8%	
25 結合動作	183	20%	1%	23%	5%		10%	2%	1%	51%
26 生成	276	24%	0%	13%	1%	1%	9%	3%		
27 消滅・破壊	96	22%	39%	2%	8%			6%	1%	41%
28 精神的行動	0									
29 精神的移動	1,023	14%	4%	25%	17%	1%	42%	8%	6%	22%
30 知覚動作	194	10%	2%	30%	20%	2%	16%	18%	1%	32%
31 感情動作	1,341	20%	18%	21%	18%	1%	17%	15%	8%	13%
32 思考動作	1,374	24%	8%	20%	7%	1%	25%	10%	2%	27%
33 使役	16	6%	6%	6%	44%		63%	13%	44%	6%
34 可能	11	27%		36%						36%
35 開始	99	19%	2%	3%	1%	1%	31%	12%	3%	48%
36 終了	100	24%	28%	1%	4%				1%	52%

※ 用言意味属性ごとのパターン集における割合を表す。空欄は、全く対応付けられなかったことを表し、0%は、1%未満であることを表す。

- [19 所有的移動]は、《喜び》の原因となりやすい。対応付けられたパターンを詳しく見ると次のことがわかった。

(ア) 「(称号、褒賞、資本・金銭等を)与える」のように、〈獲得〉だけでなく〈賞賛〉や〈厚遇〉とも対応することがある。つまり、用言意味属性の文字通りの意味の他に、ある状況が加わることにより情緒の原因としての解釈が広がることがある。

(イ) 「(資材を海外に)仰ぐ」のように、物の獲得が未だ生じておらず未来に所有的移動が起こりうることを表す用言が、[19 所有的移動]の意味属性を持つことがある。その場合〈協力〉による《期待》が対応付けられやすい。また、「助成する」のように、所有的移動の起こった後に、連鎖して事象の生じることが当

然である用言もある。このときも同様に〈協力〉による《期待》が対応付けられやすい。

- [27 消滅・破壊]は、逆に、〈喪失〉による《悲しみ》が多く対応すると予想される。また、《喜び》は対極となる情緒であるので逆に対応は少ないと予想される。そこで、詳細を見てみると次のことがわかった。

(ウ)「M1 が壊れる」のように、情緒主が結合価パターンに含まれていない場合、対応付けを行う条件が満たされない。したがって、対応付けの割合が必ずしも 100%になるわけではない。つまり、所有格の人物が情緒主となりうることへの配慮が必要である。

(エ)「根絶する」や「撲滅する」のように、無くすることを能動的に実施することを表す用言は、動作主に意図があると解釈される。さらに、1つや2つが無くなっても「根絶する」や「撲滅する」という訳ではなく、多くのものを無くする動作の労力が思い浮かぶ。意図性や労力から、〈完遂〉による《喜び》が対応付けられたと考えることができる。つまり、意味属性に加えて、用言の表す語義の特徴的な部分に注目して生起特徴が対応付けられている。

以上のことより次のことが言える。

- **アスペクトの問題:** 現在の日本語語彙大系での用言意味属性は、用言の現在時制での状況の他に、未来の時点での状況を表すことがあるため、時間軸を意識して複数の用言意味属性を付随させることで、情緒推定に関してのスムーズングに効果があると予想される。
- **隠れた情緒主をパターン上に明示させる問題:** 日本語語彙大系では日英機械翻訳のための知識ベースであるため問題にならなかったが、情緒推定では「所有格」の人物が情緒主となりうることもある。その他に、行為の同伴者が挙げられる。情緒推定に向けては、日英機械翻訳では任意であった部分でも変数で明示しなければならない。
- **用言ごとの特殊性を記述する問題:** (ア)や(エ)のように、用言意味属性ではカバーできないような用言の語義の特殊性を表すために、どのような記述を結合価パターンに追加すればよいのであるか。(ア)のように「与える」という用言は格要素に体言意味属性を制約条件として記述すれば解決できることもあるが、(エ)のように新たな中間的な記述が必要なこともある。情緒推定のためのスムーズングに向けては、今後、記述を検討する必要がある。

4.3 日本語の用言と情緒の曖昧性

言語表現の解析全般において曖昧性が常に問題になる。情緒推定においても例外ではなく、用言に付随する情緒も一般に曖昧である。用言に付随する情緒の曖昧性を解消するためには、用言の語義を区別すること、情緒主を区別すること、および、動作対象と情緒主との関係を考慮することの3つが必要である。

たとえば、多義の多いことで知られている「掛ける」は、「攻撃を掛ける」の場合、攻撃される者(情緒主)に《悲しみ》や《怒り》が生起すると推定され、「目を掛ける」の場合、掛けられた者(情緒主)に《好ましい》が生起すると推定される。また、情緒主の区別については、「略奪する」の場合、略奪した人は〈獲得〉による《喜び》が、略奪された人は《怒り》や《悲しみ》が生起する。さらに、情緒主と動作対象などの関係については、「入手する」という用言においては、入手した者(情緒主)にとって必要な物を入手したのであれば《喜び》が生起するが、つまらない物を入手したのであればその生起はない。

定性的には上述のとおり曖昧性を解消する手だてが考えられる。そこで、本節では、用言の語義に対応する情緒の曖昧性を定量的に調査する。

4.3.1 情緒の曖昧性のある用言

まず、語義の区別と情緒主の区別をせず、用言ごとに対応付けられた情緒の種類数より、曖昧性を調べる^{*1}。結果を表9にまとめる。日本語語彙大系の全用言 6,118 個に対して、情緒《なし》も含めて曖昧性のない用言は、64.5%であり、曖昧性のある用言は、35.5%であった。全体的に見れば曖昧性のある用言の割合は低い。しかし、曖昧性の高い用言をみると、和語の動詞であった(表10)。和語動詞は使用頻度が高いため、実際の情緒推定においては曖昧性の問題が無視できないと予想される。

4.3.2 曖昧性解消の可能性

語義の区別と情緒主の区別が情緒の曖昧性解消にどの程度効果があるかを調べた結果を表11にまとめる。表内の数値は、用言の数である。「語義のみ区別」の列は、一つの用言に複数の結合価パターンが存在するとき、各

*1 一般的な比較をするために、用言の語義(結合価パターン)の区別をあえてしていない。

表9 用言に対応付けられた情緒の曖昧性

情緒の曖昧さ (《なし》を含む情緒の種類数)	割合	該当する用言 (該当数 / 総用言数)
9	0.0%	(3 / 6,118)
8	0.3%	(16 / 6,118)
7	0.4%	(22 / 6,118)
6	0.5%	(31 / 6,118)
5	1.1%	(70 / 6,118)
4	3.0%	(184 / 6,118)
3	8.2%	(504 / 6,118)
2	22.0%	(1,344 / 6,118)
1	64.5%	(3,944 / 6,118)

表10 曖昧性の高い用言の例

情緒の曖昧さ	用言の例
9	掛ける, 打つ, 付く
8	する, 起こす, 見る, 使う, 取る, 受ける, 出す, 上がる, 成る, 当たる, 踏む, 突く, 入れる, 付ける, 立つ
7	ある, 引く, 飲む, 下す, 開ける, 掛かる, 遣る, 言う, 守る, 挙げる, 乗る, 折る, 掻く, 置く, 張る, 転じる, 当てる, 抜く, 変える, 与える, 落とす, 濡らす
6	逸らす, 引っ張る, 押す, 外す, 呼ぶ, 固める, 絞る, 行く, 差す, 止める, 示す, 収容する, 集める, 食う, 振る, 切る, 接する, 潰す, 吐く, 塗る, 買う, 払う, 変わる, 返す, 報いる, 無い, 明かす, 利かせる, 立てる, 流す, 裂く

結合価パターンごとに対応付けられた情緒の種類数の最大値を、その用言の「情緒の曖昧さ」とした。「情緒主のみ区別」の列も同様に、1つの用言において複数の情緒主が存在するとき、各情緒主ごとの情緒数を用いた。和語と漢語により多義の度合いが異なるため、区別してまとめた。語義や情緒主の区別をしない場合(“区別なし”の列)は表9と同一である。表11より次のことが分かる:

- 和語は、語義の区別が情緒の曖昧性解消に効果があり、情緒主の区別はそれほどではない。
- 漢語は、情緒主の区別のほうが、語義の区別よりも効果が若干高い。
- 語義の区別と情緒主の区別は両方を行うとより効果がある。
- 両方の区別においても曖昧さが残された部分は、情緒主と動作対象などとの関係により曖昧性解消を行う必要がある。その割合は、20.6% (1,262 / 6,118) である。

表11 語義の区別と情緒主の区別による曖昧性解消

情緒の曖昧さ	区別なし		語義のみ区別		情緒主のみ区別		両方を区別	
	和語	漢語	和語	漢語	和語	漢語	和語	漢語
9	3	0	0	0	0	0	0	0
8	16	0	0	0	1	0	0	0
7	22	0	0	0	8	0	0	0
6	29	2	0	0	15	0	0	0
5	64	6	0	2	18	1	0	0
4	143	41	27	21	50	17	6	13
3	309	195	150	142	129	53	45	49
2	626	718	574	611	534	596	565	584
1	1,974	1,970	2,435	2,156	2,431	2,265	2,570	2,286

4.3.3 対極的な情緒の曖昧性

情緒推定は元来曖昧性が完全に解消できるものではないかもしれない。しかし、情緒推定を応用する上で問題になる曖昧性は、推定された情緒を、《Positive》, 《Negative》, 《なし》に大別したとき、対極となる《Positive》と

《Negative》に分かれる場合である。そこで、本稿では9分類系の情緒として《喜び》, 《好ましい》, 《期待》, 《悲しみ》, 《恐れ》, 《怒り》, 《嫌だ》, 《驚き》, 《なし》を用いていたが, 《Positive》, 《Negative》, 《なし》の3分類系に変換し, 対極となる曖昧性が, どの程度存在するのかを調べる。《Positive》は《喜び》, 《好ましい》, 《期待》と, 《Negative》は《悲しみ》, 《恐れ》, 《怒り》, 《嫌だ》と, そして, 3分類系の《なし》は《驚き》と9分類系の《なし》とをそれぞれ対応させた。表12に, 語義の区別と情緒主の区別をしない場合とした場合について, 対極的な曖昧性のある用言がどの程度存在するのかをまとめた。この表より次のことがいえる:

- 語義・情緒主の区別により, 対極的でない曖昧性が大きく解消された。しかし, 対極的な曖昧性のある用言は, それだけでは解消されない。
- 語義・情緒主の区別した用言の12.2%(748 / 6,118)が対極的な曖昧性を持つ。該当する用言の例を表13に示す。用言意味属性ごとにみて該当パターン数の多い順に例示する。このような用言は, 特に, 情緒主と動作対象などの関係に対する判定条件が必要である。

以上より, 情緒推定タスクにおいて対極的な情緒の曖昧性について, 語義および情緒主を区別しても解消できない用言は, 日本語語彙大系の用言の中で, 12.2%であることがわかった。その解決は, 前提条件の作成が必要である。今後の条件作成では, まずはこれら748個の用言に重点を置くべきといえる。

表12 対極的な曖昧性のある用言の数

語義・情緒主の区別の有無	曖昧性がない	対極的でない曖昧性がある	対極的な曖昧性がある
なし	4,359	640	1,119
あり	5,341	29	748

表13 対極的な情緒の曖昧性のある用言(一部)

パターン数	用言意味属性	用言の例
259	属性変化	魔法に掛ける, (組織へ)乗り出す, 脱却する, 任命する, 鍛える, 転職する, 増減する, 火に油を注ぐ, 壊す, …
178	身体動作	出社する, 逃げる, (たばこ/薬品を)吸う, 随行する, 歩く, 登山する, (怪我を)浸す, 口にする, 伸ばす, 処刑する, …
166	思考動作	脳味噌を絞る, 思案する, 取り組む, 熟慮する, 選出する, 認める, 例に挙げる, 探る, 腹を決める, 自粛する, …
154	精神的移動	曝す, 物語る, 頭を突っ込む, (反対を)叫ぶ, (命令が)下る, 説く, インタビューする, 授業する, 伝染する, …
135	属性	休学する, 遠ざかる, 代表する, 目立つ, 相殺する, 一変する, 堅持する, 酷似する, (願望・失望と)合う, …
108	感情動作	失恋する, 気が高ぶる, 追従する, 気に掛かる, 差別する, カンニングする, 引く, 過保護にする, 胸を膨らむ(せる), …
106	物理的移動	転じる, 出る(ていく), 吹き払う, (地域・陸地に)渡る, 遠征する, 送還する, 肉薄する, 移住する, 移民する, …
97	相対関係	水を開ける, 顔を突き合わせる, (スポーツで)タイに成る, 対決する, 嗜好が違う, 統率する, 相当だ, 仕える, …

※ 用言に対して1個以上のパターン・意味属性が存在するため, 本表のパターン数の合計は, 748を超える。

5 考察

5.1 原因に着目した情緒推定の可能性

本稿は, 情緒生起の原因を表す日本語の用言の分布を調査することが主たる目的であるが, それは, 今後, 原因の表現を手がかりとした情緒推定を実現することに備えるためのものである。対応付けを行った結果はそのための知識ベースとして利用することを念頭に置いている。そこで, 本節では, 本推定方式について, 適用できる割合と精度について実験を行う。なお, 実験には情緒タグ付き日記コーパスを使用するので, 実験に先立ち説明する。

5.1.1 情緒タグ付き日記コーパス

実験には, 日記文を使用する。日記文は, 現在, インターネット上のblogで多く見られるテキスト形態であり, その

中から流行や評判の抽出が行われることから、分析価値の高いテキスト形態である。本稿で参照する日記は、石原の図書[22]の日本語部分を使用する。この図書には、情感で章立てされて日記のサンプルが挙げられているので、情緒の種類をカバーするように実験ができる。

この日記に対して、情緒タギングの経験者6名が日記の書き手の情緒を表すタグを付与した。彼らの多数決により順位付きで、推定されるべき正解の情緒を決定することができる。このタグにより人間が日記を読む際に書き手の情緒をどのように推定しているのかが分かる。また、比較のために情緒タギングの初心者1名もタグ付与を行った。

表14に日記コーパスの一部を示す。#は通番、文は主に接続詞や連用中止を手がかりに分割した節(単位節と呼ぶ)である。「喜び」から「なし」までは9つの情緒名であり、アルファベットのa～gは、情緒タグを付与した人物名である。a～fがタギングの経験者でgが初心者である。7名のタグ付与は、前後の文脈も考慮しながら、独立に行った。1つの単位節に対して1個以上のタグを付与した。

コーパスの規模について、本図書に掲載されている日記のストーリーは268話であり、全文を単位節に分けると1,642文となった。そして、タグ数は13,165個となった。

表14 情緒タグ付き日記コーパスの例 (日記は[22]より引用)

#	文	喜び	好ましい	期待	悲しみ	恐れ	嫌だ	怒り	驚き	なし
241	今日、新しいめがねを注文した。			bcdefg						a
242	1週間かそこらで出来上がる。		g	ace						bdf
243	もっとオシャレな雰囲気になれるのが、	a	afg	abcde						
244	今から楽しみだ。	ae	abf	abcdeg						

5.1.2 実験1：原因に着目した情緒推定方式のカバレッジ

目的と方法

実験1では、本方式の適用可能な文の割合を求めることを目的とする。日記から、ランダムでストーリーを抽出し、単位節ごとに本方式で情緒推定の行うことのできる文かどうかを手作業で判別する。判別の基準は、次の3点である：

- 情緒主(日記の書き手)が文に登場すること(ゼロ代名詞として登場しても良い)
- モダリティ的な表現で情緒の解釈が著しく変化しないこと
- 情緒状態を表す文ではないこと、情緒的な反応を表す文ではないこと

たとえば、表14において、#241および#243は、「対象内」であるが、#242および#244は「対象外」である。#242は日記の書き手が登場せず、#244は、情緒の直接表現であるからである。

結果

ランダム抽出した単位文は540文であり、その内訳を表15に示す。対象外となった文はより詳しく分類した。この結果より、情緒主が登場し、かつ、原因に着目した情緒推定の行える箇所は、相対的に見て多いことが分かった。

表15 日記文の内訳

判別	対象外の理由	内訳 (該当 / 抽出数)
対象内	-	34% (186 文 / 540 文)
対象外	文に書き手が登場しない	28% (149 文 / 540 文)
	文が直接表現である	18% (96 文 / 540 文)
	文のモダリティが情緒的解釈を著しく変える	9% (51 文 / 540 文)
	文がダ文である	3% (17 文 / 540 文)
	その他	8% (41 文 / 540 文)

5.1.3 実験2：原因に着目した情緒推定方式の精度

目的と方法

実験2では、第3章で作成した情緒属性セット付きの結合価パターン集を用いた情緒推定の精度の評価を目的とする。結合価パターンを適切に用いる能力については本稿の狙いではないので、手作業によりパターンを照合して、文から情緒を推定する。しかし、1つのパターンに複数のセットが対応付けられているが、その使い分けについての知識データ(たとえば、前提条件)は十分に作り込めていないので、情緒主の区別は行うものの、複数の情緒を推定結果として出力する。

推定結果の評価は、正解情緒と出力情緒の一致率を用いる。正解の情緒は、情緒タギング経験者6名による多数決により上位2位かつ過半数(4名)の支持のある情緒とする。一致率の計算は次の式に基づく：

$$\langle \text{一致率} \rangle = \frac{\langle \text{正解と出力の一致した情緒の数} \rangle}{\langle \text{正解情緒の数} \rangle + \langle \text{出力情緒の数} \rangle} \times 100 (\%)$$

情緒推定の様子

日記の一例を以下に示す。

- 1) 蒲郡のおばさんからみかんが1箱届いた。
- 2) とても甘くて
- 3) おいしかった。

文1)は、「みかんが、おばさんから、(日記の書き手へ)、届く」という文であるので、日記の作者が「届く」という用言の格関係において登場している。文2), 3)は「(みかんが)甘い」、「(みかんが)おいしい」という文であるので、日記の作者は用言の格関係に登場しない。したがって、この例では文1)のみが情緒推定の対象である。

文1)について、正解の作成に携わった6名は、この文に対して、《喜び》に4名、《好ましい》に2名、《なし》に1名という付与であった(複数回答あり)。したがって、「正解の情緒」は、「《喜び》」とした。一方、「届く」の結合価パターンを参照すると、「情緒名:《喜び》, 原因:《獲得》, 情緒主: N2, 情緒対象: N1」というセットがあるため、「出力の情緒」は《喜び》となり、推定は成功した。

結果

実験の結果を表16に示す。情緒推定の対象となった文は、186文であり、正解の情緒の定まったものは126文である。正解の情緒の定まらなかったものとは、多数決において過半数の一致の無かった文である。人間による判定でさえ不安定な部分は、一般性の乏しい判定といえるため、本稿の一般性のあるものを収録したパターン集においては対象外となる。

さて、正解の情緒は、132個である。本稿で作成したパターン集による情緒推定の精度は、初心者であるg者よりも3個劣る結果であった。参考として、正解を定めたa者からf者について的一致率を求めた。これらの6名については、自身を除く5名による多数決で正解を定め、自身の推定結果を比較することで一致率を計算した。そのため正解数に違いがある。初心者と経験者に開きが見られた。

表16 情緒推定の精度

推定者	一致率	(一致数, 出力数, 正解数)
a者	67%	(72, 118, 98)
b者	65%	(66, 102, 100)
c者	76%	(74, 108, 87)
d者	66%	(66, 105, 96)
e者	68%	(70, 115, 90)
f者	67%	(70, 105, 105)
g者	50%	(66, 132, 132)
パターン集	45%	(63, 147, 132)

誤り分析

パターン集を用いた情緒推定の誤りを分析する。147個の出力のうち63個が正解となり、84個が不正解となった。84個の推定元の74文について誤り理由を考察する。結果を、誤りの種類の観点から以下のようにまとめる。

- 少数に支持される情緒と一致: 14個
- 用言の問題
 - 結合価パターンが不足: 13個
 - 情緒属性セットの改良が必要: 5個
 - 個人に依存した判断が必要: 6個
- 格要素などの問題
 - 時詞・時制による判断が必要: 9個
 - 副詞による判断が必要: 3個
 - 形容詞による判断が必要: 1個

- 内容による判断が必要: 20 個
- 構文構造による判断が必要: 2 個
- 文脈による判断が必要: 3 個
- 直接表現と原因の表現が混在: 1 個

主な物について具体例を示しながら説明する。

(事例 1: 個人に依存した判断が必要な場合)

次の事例では、個人の嗜好に依存する。今後、個人の嗜好を表す体言のリストを作成し、パターンごとに運用規則を付随させ、すなわち、「N1 が N2 に参加する」というパターンについては N2 を介して、個人の嗜好に依存した判断をする必要がある。

- 1189: 昨日、町民体育祭で障害物競走に参加した。
a 者～f 者による推定: 《喜び: 1 票》, 《好ましい: 1 票》, 《なし: 4 票》
パターン集による推定: 《好ましい》, 《嫌だ》
- 1190: とても楽しかったけど、
- 1191: 今日は全身が筋肉痛。

(事例 2: 時詞・時制による判断が必要な場合)

次の事例では、時詞や時制に依存する。未来における話し手の行動や状態が述べられる際、好都合な内容かどうかを判断し、《期待》や《恐れ》の情緒を推定する必要がある。「今年の抱負」のように一般に未来における好都合なことを連想させる名詞句(時詞)は、重要な手がかりとなる。

- 119: 今年の抱負は、英語で日記をつけることと、
a 者～f 者による推定: 《期待: 4 票》, 《なし: 2 票》, 《好ましい: 1 票》
パターン集による推定: 《なし》
- 120: ラジオで NHK の英語番組を聴くこと。
- 121: 来年までに一人でアメリカへ行けるようになっていたいなあ。

(事例 3: 内容による判断が必要な場合)

次の事例では、内容に依存して情緒が決まる。しかし、パターン集による推定では、〈判明〉による《喜び》という情緒が推定される。〈判明〉は、「目標を実現するために知る必要のあることを知ること」であるので、「ファスナーが開いていること」を知るが目標実現に役立つことかどうかを判断しなければならない。そのためには、情緒属性の中でも「前提条件」を用いる必要がある。

- 499: 帰宅したら
- 500: ズボンのファスナーが開いていたことに気づき、
a 者～f 者による推定: 《嫌だ: 4 票》, 《驚き: 3 票》, 《悲しみ: 1 票》, 《恐れ: 1 票》
パターン集による推定: 《喜び》
- 501: 恥ずかしさのあまり赤面してしまった。
- 502: 誰にも見られていなければいいけど…。

(事例 4: 構文構造による判断が必要な場合)

次の事例では、「～のに～だ。」という構造に着目して判断する必要がある。「『当然のことを実行している』のに『好都合なことにならない』」という構造の場合、文全体で《怒り》の最上位特徴を表すといえる。今後、情緒生起の原因を表す複文の構造について事例を収集する必要がある。なお、構造全体で複雑な事態を表現していることから、このような構造に着目する情緒推定は、第 2 章の分類ではタイプ 2 に該当する。

- 995: ずっと一生懸命働いているのに、
a 者～f 者による推定: 《怒り: 4 票》, 《嫌だ: 3 票》, 《なし: 1 票》
パターン集による推定: 《なし》
- 996: いまだに昇給なし。

以上のように、情緒生起の原因に着目した情緒推定において、本パターン集の限界について確認することができ

た。具体的には、個人の嗜好性や時制、構文構造などが新たに着目しなければならないことがわかった。また、本パターン集のパターンが不足することや情緒属性に改良すべき点のあることも確認できた。

今後、本パターン集による情緒推定実験を繰り返し、パターンの追加や情緒属性の改良を行うとともに、新たに着目すべき事項について検討し、パターンを用いる方法との融合を図る必要がある。

5.2 言語表現に明示されない事態・暗示される事態

表 15 によると、「情緒生起の原因を表し、かつ、日記の書き手が登場する文」が、34%であったが、その次に多いのは、「日記の書き手が登場しない文」であった。それは、たとえば、「昨日はとても良い天気だった。」や「大切なグラスが壊れた。」のような状態を説明する文や人物の登場しない用言である。これらの文の出現する割合が少なくないため、原因に着目した情緒推定(表 1 におけるタイプ 2(1))を充実させる上では無視できない種類の文である。しかし、これらの文を扱う上で問題点がある。それは、これらの文では推定しようとしている情緒の主が、文に表される事態に対してどのような関係にあるのかが明示されないため、1つの文からは判断できないという点である。

結合価パターンでは、一般に格要素の変数 N に対する所有者などは変数化されないため知識ベース上で捉えることができない。たとえば、「 $M1$ が壊れる」というパターンはあっても「 $M1$ の $M2$ が壊れる」というパターンは存在しない。もし、後者のパターンが存在するならば「情緒主: $M1$ が、 $M2$ を失うことにより悲しむ」という情緒属性をパターンに対応付けることができる。

所有者と物体の関係に限らず、事態間の関係も言語表現の背後に存在する。たとえば、「 A 氏が B 氏に X を密告する」という事態には、「 C 氏が X をする」という事態が背後に存在する。このように背後に存在する事態を考慮することで、この「密告する」に対して「情緒主 C 氏が A 氏に対して、暴露による怒りを持つ」という情緒属性を対応付けることができる。

一般に、結合価パターンのそれぞれに対して、背後に存在する事態を付随させるという分析作業は、漠然としており作業の目標が定めにくい。しかし、本稿で用いた「情緒生起の特徴」と関連付けることを目標に作業を行うことは可能である。たとえば、「密告する」においては《怒り》の最下位特徴である《暴露》が常識的に思い浮かぶため、それを具体的に表現することを目標として事態を関連付けられたい。

情緒生起の原因に着目した情緒推定を深める上で、今後、このような表現の背後に存在する事態を追跡する機構が必要になる。

6 おわりに

本稿では、言語表現から情緒を推定するタスクを 4 つに大別した。中でも情緒の生起する原因に着目して推定を行う方式のための言語知識に関する議論を深めた。具体的には、情緒の生起する一般的な原因を表す日本語の用言の分布について調査した。その分布を知るためには、日本語の用言の体系と情緒生起の原因の体系との間の対応関係を分析する必要があった。用言の体系は、[10]の「日本語語彙大系」に収録されている「結合価パターン」を参照し、情緒生起の原因の体系は、[11]、[12]の「情緒生起の特徴フレーム」に示される「特徴ラベル」を参照した。全ての結合価パターンを人手により分析し、両者の対応付けを行った。その結果、全ての結合価パターン(14,812件)の52.2%が、情緒主、および、情緒生起の原因を表すことが分かった。日本語の用言と情緒の関係に関する基礎データとして次ぎの3点を調査した:

- **情緒生起の原因ごとの用言の分布:** 情緒生起の原因を表す用言は、《好ましい》、《喜び》、《嫌だ》、《期待》を表すものが多く、《怒り》、《驚き》を表すものが少ないことが分かった。さらに、たとえば、《制限》による《嫌だ》を表す用言が最も多いことなど、細かな単位で分布を知ることができた。
- **用言意味属性ごとの用言の分布:** 用言意味属性が同一の用言を比較した。その結果、用言に関わる情緒の一致する確率を見ると、20%未満のものが多くということが分かった。このことは、情緒推定用の言語知識ベースを、用言意味属性だけでスムージングすることが、あまり良い結果をもたらさないことを示している。用言の語義の微妙な違いを区別することが重要で、特に「用言意味属性の成立するアスペクト」、「パターン上に明示されない人物」、および、「動作主の意図や目標」を扱う必要がある。
- **用言からの情緒推定の曖昧性:** 語義と情緒主の区別をせず用言のみで情緒を推定する場合、曖昧性のある用言が35.5%であるが、語義と情緒主の区別をすると曖昧性のある用言は20.6%であり、さらに対極的な曖昧性のある用言は12.2%であることが分かった。さらなる曖昧性解消には、情緒主の個別事情を考慮する必要がある。

次に、今後の情緒推定機構の実現に向けて考察を行った。情緒生起の原因に着目して情緒を推定する方式を日記に対して適用したところ、本稿で対象とした日記においては、情緒を明示する箇所が18%であることに對して、原因に着目して情緒を推定する箇所が34%であることから、原因に着目する方式の必要性が確認された。また、本

稿で対応付けたパターンを手動で適用して情緒を推定した結果、情緒推定の初心者が行った情緒推定の精度と同等であることが確認された。その誤り事例をみると、時制、副詞、文脈、個人的事情などの考慮が必要であるが、基本的な部分としては、原因に着目した情緒推定が正しく行えていたといえる。今後、情緒推定の機構を構築する上で考慮すべき点が具体例とともに確認できた。

なお、結合価パターンと特徴ラベルの対応付け作業に対する評価は、作業の途中段階で行った検査結果によると誤り率が14.2%でありそれらは修正済みであること、および、試行的な情緒推定実験の結果より一定レベルの信頼性(情緒推定の初心者程度)が確保できていると言える。また、分析の基盤とした「情緒生起の特徴フレーム」は、最下位特徴が用言に対応付けられる割合が93.2%と高く、「その他の特徴」の割合が低いことから、特徴ラベルの網羅性が確認できた。一方、「日本語語彙大系」は、情緒推定実験において結合価パターンの不足する件数が7%であったことより網羅性が確認できた。

以上のように、本稿では、情緒生起の原因を表す用言という新しい観点から日本語の用言を分析できた。今後、情緒推定を目指して語彙分析を深めることで、用言に暗示される事態を明らかにしたい。

謝辞

本研究は、科学研究費補助金(若手研究 B:17700151)の支援の下で行いました。対応付けに対する検査作業に協力した鳥取大学大学院生の石上真理子さん、岡本一輝さん、および、金澤佑哉さんに深く感謝します。また、情緒タグ付き日記コーパスの構築に協力して頂いた研究室メンバーに感謝します。

参考文献

- [1] 飛田良文, 浅田秀子: 現代形容詞用法辞典, 東京堂出版, 1991.
- [2] 飛田良文, 浅田秀子: 現代副詞用法辞典, 東京堂出版, 1994.
- [3] 飛田良文, 浅田秀子: 現代擬音語擬態語用法辞典, 東京堂出版, 2002.
- [4] 大野晋, 浜西正人: 角川類語新辞典, 角川書店, 2002.
- [5] グループ・ジャマシイ: 日本語文型辞典, くろしお出版, 1998.
- [6] 金子真: 焦点化副詞「ナンカ」が表わす否定的評価の派生について, 言語処理学会第12回年次大会ワークショップ「感情・評価・態度と言語」論文集, pp.33-36, 2006.
- [7] 目良和也, 市村匠, 相沢輝昭, 山下利之: 語の好感度に基づく自然言語発話からの情緒生起手法, 人工知能学会論文誌, Vol.17, No.3, pp.186-195, 2002.
- [8] 遠藤大介, 齋藤真実, 山本和英: 係り受け関係を利用した感情生起表現の抽出, 言語処理学会第12回年次大会発表論文集, pp.947-950, 2006.
- [9] 土屋誠司, 吉村枝里子, 渡部広一, 河岡司: 状況表現を考慮する発話文章からの感情判断手法の提案, FIT2006 情報技術フォーラム, 一般講演論文集, 第2分冊, pp.231-234, 2006.
- [10] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系, 岩波書店, 1997.
- [11] 徳久雅人, 岡田直之: 情緒の生起過程へのパターン理解の接近, 電子情報通信学会技術研究報告, TL95-8, pp.9-20, 1995.
- [12] 徳久雅人, 岡田直之: パターン理解的手法に基づく知能エージェントの情緒生起, 情報処理学会論文誌, Vol.39, No.8, pp.2440-2451, 1998.
- [13] 南不二男: 現代日本語文法の輪郭, 大修館書店, 1993.
- [14] ダニエル ヴェンダーヴェーケン: 意味と発話行為, ひつじ書房, 1997.
- [15] Fiorella de Rosis, Floriana Grasso: Affective Natural Language Generation, Ana Paiva(Ed.), *Affective Interactions*, LNAI 1814, pp.204-218, Springer, 2000.
- [16] Masato Tokuhisa, Ryoko Tokuhisa, Kentaro Inui, Naoyuki Okada: Emotion Recognition in Dialogue, G.Hatano, N.Okada, H.Tanabe (Eds.), *Affective Minds*, pp.221-229, Elsevier, 2000.
- [17] 古塩貴行, 徳久雅人, 村上仁一, 池原悟: 情緒注釈付きコーパスの誤り分析, 人工知能学会全国大会, 2G3-02, 2004.
- [18] 往住彰文: 認知的感情の構造と文学テキスト理解, *Cognitive Studies*, Vol.8, No.4, pp.369-375, 2001.
- [19] 松本和幸, 任福継, 黒岩眞吾: 語の意味情報を考慮した感情推定アルゴリズム, 言語処理学会第11回年次大会発表論文集, pp.145-148, 2005.
- [20] Andrew Ortony, Gerald L. Clore, Allan Collins: *The Cognitive Structure of Emotions*, Cambridge University Press, 1988.
- [21] Clark Elliott: The Affective Reasoner: A process model of emotions in a multi-agent system, Northwestern University, Ph.D Thesis, 1992.
- [22] 石原真弓: 英語で日記を書いてみる, 表現集編, ベレ出版, 2002.