

俳句テキストの語彙分析に基づく知識発見に向けて

鈴木 雅実[†]皆川 直凡[‡]KDDI研究所[†]鳴門教育大学[‡]

1 はじめに

俳句は、五七五の有季定型で凝縮された作者の想いを伝える、日本発の CGM (Consumer Generated Media) の先駆けであり、感性コミュニケーションメディアと見做すこともできる (鈴木 2006)。本研究では季語を含む語彙の組合せの特徴を計量的に分析することにより、多くの共感者を得る俳句/そうでない句、初心者/熟達者の句、散文との比較結果等から得られる、新たな知識発見を目指す。今回は創造性や感性面での探求も視野に入れて、最初に分析対象とする俳句コーパスに基づく研究の着眼点とねらいについて述べる。

2 目的と背景

従来、俳文学等の人文科学系の研究分野では、松尾芭蕉をはじめとする著名な俳人が詠んだ句の特徴を比較分析することは一般的であったと思われるが、必ずしも定量的な考察ではなかったかもしれない。一方、俳句テキストを対象とした情報科学的な研究としては、句の中に詠み込まれた鐘の音や虫の声等の語の分布状況 (場所や時間等の他の語句との取合せを含む) から、各時代ごとのサウンドスケープ (音環境) を推定する等の先行研究が見られる (鈴木 2006 の中で紹介)。

皆川は、自身の先行研究 (皆川 2005) に基づいて選定した 3 因子計 15 の形容詞対から成る Semantic Differential 尺度を用いて、教科書俳句の情緒的意味と既知度の評定を求めた。その結果、俳句への興味・関心の程度により、評定傾向が異なることが明らかとなった (皆川 2008)。

本研究は、心理言語学的ないし教育心理学的な視点を含む学際的な見地から、俳句という表現スタイルに凝縮された語の組合せが持つ特徴を様々な観点から分析することにより、表現力や創造力

Towards Extracting Knowledge based on Haiku Text Analysis

[†] Masami Suzuki, KDDI R&D Laboratories Inc.[‡] Naohiro Minagawa, Naruto University of Education

に関係する要因の科学的な解明に結びつくような知識発見を目指すものである。

3 対象とする俳句テキスト

四国遍路についての 5 回の事前授業と 2 泊 3 日の歩き遍路 (7~12 の寺院を巡る) によって構成される大学および大学院 (受講生には、小・中・高校の現職教員を含む) の授業において、歩き遍路の途上で出会う風物や、同行者や地域の人々とのふれあいをテーマとして俳句を作ることを提案し、受講生に俳句の提出を求めた。第 2 著者が事前授業を担当した 1 回の授業では、まず、俳句の面白さ、豊かさは、有季定型によって生み出されることを伝え、その仕組み (十七音、季語) について、解説した。そして、「遍路の途上では、さまざまな風物に出会う。句の食材にも出会う。その一つ一つが季語なのである。歩き遍路は俳句をつくる絶好の機会であるといえる。この機会に、自分自身の心のかたちを、俳句という五七五の十七音に書き換えていく楽しさを味わってみよう。」と呼びかけた。さらに、以下のように、俳句の創作法を教授した。

「俳句は十七音で、季語が一つ入っていればよい」
「俳句の基本リズムは、五/七・五 or 五・七/五」

手順① 俳句に詠みたい季語を見つける。

手順② 季語と直接関係はないが、イメージの近い十二文字の文を作る：

- ・十二文字 (五文字+七文字) の文+「季語 (+かな)」
 - ・「季語 (+や)」+ 十二文字 (七文字+五文字) の文
- ただし必ずしもこの方法によらなくてもよい。

歩き遍路の実施時期は、9 月下旬、残暑の候である。出発時に、俳句創作法入門、秋の季語集、メモ用紙から成る「俳句のしおり」を配付し、俳句の募集についての説明を行い、夜のミーティング時に、創作についての助言を行い、俳句の提出を求める。この取り組みは、2007 年度より 3 年間継

続しており、現在までに、239名より計717句の俳句が収集されている。

2009年10月に、20名の歩き遍路参加学生を集め、作者名を伏せて鑑賞会を開いたところ、活発な読み合いが行われた。全156句を2回に分けて音読・鑑賞し、選評を述べ合った。最後に、選評文の提出を求め、いくつかの俳句が複数の読み手の共感を得て、選句された。その例を以下に示す。選句者の数を括弧内に示す。

空海と初秋の風に背を押され (8)
知らぬ間に澄んだ心と秋の空 (8)
杖のへり共に歩んだ秋遍路 (8)
秋遍路上り下りは人生路 (6)
先人の道を踏みしめ秋の山 (6)
秋の雨 滴と遊ぶ彼岸花 (6)
遍路道花咲く笑顔と秋桜と (5)
草の花ふまれてもなお美しく (5)
竹林のすき間に見える秋の空 (4)
目的地近づくたびに秋気澄む (3)

共選句の多くは、情景描写が適確で具体性があり、作者の心情が適度に表現され、想像の余地を残す俳句であった。そのことは、同じ句の選者の選評が共通部分と個別部分から成ることにより推察された。

この俳句テキスト・コーパスの特徴と研究対象として適切と思われる点は次の通りである。

- 個々の俳句に共通した文脈の存在
遍路中のほぼ同一条件下で詠まれた句が集められており、比較のための基礎が整っていると考えられる。
- 統計的な分析を可能とする句の数
同一（または類似の）季語を含む句が一定数集まっていることから、定量的に意味のある考察をするに十分である。
- 作者の属性情報が参照可能
歩き遍路への参加者の種々の個人的な属性による傾向の違いも観察できる。

これに対して、一人のあるいは複数人の作家による句集を対象とした場合、通常は季題やテーマ等についてオムニバスの編集がなされる場合が多いことから、統計的に意味のある知見を得るにはあまり相応しくないと考えられる。

4 分析の着眼点とねらい

前述したような3年間に渡る「歩き遍路」という共通の背景の下に詠まれた俳句を対象として、季語の使われ方や作品としての水準（選句の多寡）等の観点から分類するとともに、季語を含む各語彙とその周辺の他の語彙との同時出現（共起）頻度等を観測する。これらの基礎データを基に、次のような比較分析を試みる。

- 季語別に分類し、どのような季語が用いられやすいかを探る。
- 俳句の中で季語の含まれる部分と含まれない部分とがどのような意味的關係をもっているか、それが選句者（共感者）の多寡、あるいは初心者と熟達者の違いとどのように関わっているか。
- 俳句の構成語句間の結びつきが一般の文（新聞のようなテキスト・コーパスを想定）とどのように異なるかについての数量的分析
- 意外な取合せと感じられる語彙の同時使用に関する傾向と、共感の多寡との関係

これらの比較分析結果に基づいて、俳句の創作を通じた、表現力や創造力に結びつく諸相の關係性が少しずつ明らかになることが期待される。

5 おわりに

本研究は学際的な研究による言語処理関連の新しい知識発見を目指すものであるが、分析対象として適切かつ興味深い俳句テキスト・コーパスが得られたことから、その具体化に着手した段階である。実際の分析結果から得られる知見については今後発表を予定する。なお、今回の着想に至るまでに、著者の一人（皆川）による日本心理学会での俳句関連ワークショップの連続企画と、そこでの交流が契機となり検討を開始した。合わせて参加討論頂いた関係各位に謝意を表したい。

参考文献

- 鈴木雅実 (2006) . 感性コミュニケーションメディアとしての俳句, 小特集「感性とコミュニケーション」(庄司裕子編), 人工知能学会誌, Vol.21 No.2, pp.189-194.
- 皆川直凡 (2005) . 俳句理解の心理学, 北大路書房.
- 皆川直凡 (2008) . 俳句への興味・関心が俳句の情緒的意味の評定に及ぼす影響—創作経験の少ない鑑賞者を対象とするSD法による検討—, 鳴門教育大学情報教育ジャーナル, 第5号, pp.67-70.