

共通体験を通じて詠まれた俳句からの知識発見の試み

An Attempt to Discover Knowledge from Haiku by Authors with Common Experiences

鈴木 雅実^{*1}
Masami SUZUKI

皆川 直凡^{*2}
Naohiro Minagawa

^{*1} KDDI研究所
KDDI R&D Laboratories Inc.

^{*2} 鳴門教育大学
Naruto University of Education

Haiku is one of earlier CGM originated from Japan and can be regarded as a type of "Kansei Communication Media". In this research, we launched an analysis of vocabulary distribution in Haiku corpus, which contains about 700 haiku by the participants of Autumn Pilgrimage in Shikoku, Japan. Thus, our ultimate goal is knowledge discovery from haiku based on common experiences that will be related with creativity or various human factors.

1. はじめに

本研究は、俳句テキストデータからの知識発見を目指すものであるが、対象とする俳句作品は、教員養成系の大学学部および現職教員を主とする社会人を含む大学院の学生による四国遍路に関する共通テーマを扱ったものである。すなわち、四国遍路についての5回の事前授業と2泊3日の(7~12の寺院を巡る)歩き遍路という参加者の共通体験の下で詠まれた俳句作品を分析することにより、表現力や創造力に関する要因を種々の観点から探り出そうとする意図に基づいたものである。また、個人または複数の作者によるオムニバスの句集を対象とした場合とは異なり、ほぼ類似した背景で創作された一定数の俳句から、ある程度統計的に意味のある傾向が観察されることが期待される。今回は、深い分析に入る前の予備検討として、分析対象とする俳句テキストの表層的な語彙分布等から得られた、対象コーパスの特徴と今後の分析の着眼点について述べる。

2. 背景と分析対象

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している。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、表現力や創造力に関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[鈴木2010]。

このような背景の下で、第2著者が指導した四国遍路に関する授業では、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、以下のように基本的な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」

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、今回分析の対象とする俳句は3年分の計713句である(表1参照)。

上記の分析対象句のうち2009年度分については、遍路体験後の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。その際に選句された作品を選句者数(括弧内)とともに例示する。

- 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)

3. 基礎データの収集と表層的な観察事項

3.1 年度毎の作者の属性

上記のような3年間に渡る「秋遍路」における俳句創作の参加者計240名と、詠まれた句数の分布状況は表1の通りである。

表1. 受講生(俳句提出者)の属性分布

年度	男性		女性		合計	提出句 (平均)
	学部生	大学院	学部生	大学院		
2007	0	32 * 11	0	14 * 5	46 * 16	201 (4.4)
2008	37	16 * 5	55	7 * 1	115 * 6	354 (3.1)
2009	32	8 * 2	34	5 * 2	79 * 4	158 (2.0)
合計	69	46 * 18	89	26 * 8	240 * 26	713 (3.0)

*は内数で現職教員(社会人)

3.2 俳句に含まれる語彙の分布

今後のより深い分析では、選句の多寡等や作者の属性等との比較検討を行う予定であるが、手始めとして各年度における俳句作品に含まれる語の出現頻度を調査した。年度毎の俳句からの総抽出語数および異なり語数(十名詞類)を表 2 に、また出現頻度の高い季語とそれ以外の語を表 3 に示す。なお、今回の分析にはテキスト分析ツール KH-Coder を使用した[樋口 2010]。

表 2. 年度毎の俳句に含まれる語数の基礎分布

年度	句数	総抽出語数	異なり語数(うち名詞)	名詞のうち頻度 3 以上
2007	201	1426	611 (260)	19.6%
2008	354	2730	909 (379)	18.5%
2009	158	1285	507 (210)	15.7%
全体	713	5441	1471 (630)	21.4%

表 3. 年度毎の俳句に詠まれた出現頻度の高い語

年度	季語	頻度	季語以外	頻度
2007	秋遍路	28	遍路道	33
	秋(の)風	16	歩く	14
	コスモス=秋桜	12	道	13
	彼岸花 (=曼珠沙華)	10	声	9
	<以下順に>		寺	8
	秋(の)山/虫/秋暑し		<以下順に> 心/水/輝く	
2008	秋遍路	49	遍路道	44
	曼珠沙華 (=彼岸花)	40	道	28
	秋(の)風	27	歩く	25
	赤とんぼ=赤蜻蛉	24	心	18
	<以下順に>		見る	16
	秋(の)空/秋(の)山 秋桜=コスモス		<以下順に> 声/吹く/足	
2009	秋遍路	45	遍路道	15
	秋(の)雨	16	歩く	11
	彼岸花	10	道	11
	秋(の)空	9	心	9
	<以下順に>		<以下順に>	
	霧/秋(の)風		水/声/見る	

表 2 および表 3 から、年度毎の俳句に詠まれた語の分布状況は、ほぼ類似していることが分かる。ただし、各年度の天候の違いや目にすることの多かった自然の風物等は、高頻度で使用された語彙に反映されている。例えば 2007 年は残暑が厳しく、2008 年度は赤とんぼの類が多く観察されたこと、さらに 2009 年度は雨模様で花の盛りを過ぎていたことが如実に現れた。さらに、頻度 2 以下の名詞が 8 割程度を占めることから、作者が注目した対象は相当多岐に渡っていることが窺え、共通体験の下での俳句作品の多様性の源とも考えられる。なお、上記の表 3 に挙げた季語については、「秋(の)空」=「秋空」のように助詞が入るなどの表記が異なるものも合わせて一つの単位として認定した。

ここで、最頻出語と言える「遍路」の使われ方に着目すると、表 3 に示すように「秋遍路」という季語の形または「遍路道」のような表現として、句に読み込まれている。「遍路」だけでは本来は春の季語とされ、秋+遍路により秋の季語となることは参加者に予め教示された。次に示すのは、「秋遍路」と同時に使用された語の分布である。今回のデータはいずれも絶対頻度を挙げている。

表 4. 「秋遍路」と同時に詠まれた高頻度の語

年度	「秋遍路」との共起頻度の高い語(度数)
2007	歩く(5) 忘れる(3) 感じる(2) 思い・思い出/時空/接待/空海/我/陽(各 2)
2008	歩く(8) 励ます(3) 優しい(3) 着く/登る/永い/険しい(2) あいさつ/いつのぼり坂/自分/笑顔/仲間/頂上/友達(2)
2009	歩く+歩む(7) 滑る/感じる/見る/通る(2) バス/季節(2)
全体	歩く+歩む(21) 感じる(5) 忘れる/励ます/優しい(4) 滑る/見る/行く/咲く(3) 思い/自分/笑顔/仲間(3)

表 4 から、年度間の共通点ならびに相違点を見出すことができる。まず、「秋遍路」が「歩く」という語と同時に使用される頻度の多さは一貫しており、歩くことへの印象の結び付きの強さが読み取れる。年度ごとの特徴をみると、2007 年度では「接待」や「空海」、2008 年度では「仲間」や「友達」に特徴を見出すことができ、歩きながら遍路の意味を感じたり考えたりするゆとりを示した作品が多い。一方、2009 年度は「滑る」「通る」に代表されるように、歩くことへの専心が感じられる。これには、表 3 の頻出語の分布との対応でも述べたように、2007 年度と 2008 年度は概ね好天であったのに対し、2009 年度は初日が終日雨であったという気象条件の違いが強く反映されているのかもしれない。

4. 今後の分析の着眼点 ~おわりに~

前章で代表的な季語として取り上げた「秋遍路」とは対照的に、「遍路道」が使用された俳句においては、俳句作法の通り秋の季語(秋風・コスモス・赤とんぼ)などが同時に詠み込まれているほか、最頻の共起語として「心」が出現しており、「歩く」の類の動詞類はほとんど見られない点が興味深い。今後さらに注意深く多面的な分析を必要とするが、このことは「秋遍路」と「遍路道」という個別の表現から想起されるイメージ、およびそれと同時に用いる語の持つイメージの重なり/補完の関係が、俳句という表現形式で最大の効果を発揮するように、語彙の選択が行われているのではないかという洞察に繋がる可能性がある。

一方、2 章に示した高評価を獲得した句の多くは、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。この理由としては、事物の表層的な意味関係の裏に、作者の深い心情が重層的に表れされていることが読み取れるような句が選好されることが多いという知見[皆川 2005]と符号するものである。このような句の理解については、個人の体験や実世界知識等との結び付きについて想像をめぐらせる必要があり、前章のような語の出現分布だけでは説明がつかないと考えられるが、知識発見の課題としてチャレンジしたい。

最後に、今回の分析への着手については、第二著者(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ったことを記し、合わせて参加討論頂いた関係各位に感謝申し上げる。

参考文献

- [鈴木 2010] 鈴木雅実・皆川直凡: 俳句テキストの語彙分析に基づく知識発見に向けて、言語処理学会第 16 回年次大会, E4-3, 2010.
- [鈴木 2006] 鈴木雅実: 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, 2006.
- [皆川 2005] 皆川直凡: 俳句理解の心理学, 北大路書房, 2005.
- [樋口 2010] 樋口耕一: KH Coder Index Page (<http://khc.sourceforge.net/>)