

浅い解析を出発点とする俳句の形式的解釈

新 田 義 彦

I はじめに

浅い解析 (Shallow Parsing) ([1] ~ [8]) とは、文を構成する字句あるいは語句を、その並びに沿って順次走査し、簡単なラベル (語句識別子) を付与しつつ語句を分節する処理である。文の句構造や依存構造の認識や解析には立ち入らない。このような単純さのゆえに、広範囲の文形の取り扱いが可能になり、解析に失敗する事態を微小に抑えられる。対象言語の特性に対する依存性から独立できる、といったような頑健性を担保できる (Kinyon, A. (2001))。

頑健性を入手した代償は、詳細な言語的知識、文構成情報が入手できないことである。

浅い解析は、プログラミング言語の字句解析 (綴りや形式の検査)、ワードプロセッサにおける入力 (アルファベット列やひらがな列) の綴りの誤り検査と訂正、などに効果的に利用されている。

このように浅いレベルでの形式検査に適しているので、文の意味解析のような深い解析には使えないと考えられてきた。この考え方は決して間違っていないのであるが、本論文では、浅い解析処理を出発点として、俳句のような断片文の意味解釈に接近する方法について検討する。考察する方法の眼目は、語のオントロジー知識、および語と語を接続する助詞の精細な弁別である。

浅い解析を出発点とする理由は、下記のように要約できる。

- 1) 俳句のようにきわめて短い断片文には、元来、主語 + 補語、述語 + 対象、主体 + 修飾 (限定)、といったような完成した文構造が

存在しない。したがって、句構造解析や依存構造解析 ([9] ~ [14]) を行う意義が希薄である。

- 2) 俳句のごとき断片文を構成する、語句に何らかの標識 (ラベル) を付与した記号列を出力し、記号列の解釈を後続する何らかの処理に託するのが妥当である。
- 3) 後続する何らかの処理は、必然的に語句の意味あるいは語用論的效果の推定処理となる。推定には歳時記などのオントロジー知識を利用するが、これについては後の章で述べる。

II 俳句の形式的意味

俳句を形式的に解釈するとは、つまり俳句の意味を形式的に捕えることである。本論文では下記のように、「核文の抽出 (同定)」と「核文の間の意味的關係の推測 (概略的同定)」という 2 段階形式で、形式的意味を捉えている (新田義彦 (2012-3) および新田義彦 (2013-2))。

俳句文を H とする。

- 1) H から、対峙する 2 つの核文 K 1 と K 2 を抽出すること、

そして

- 2) 核間関係 R を導出すること。

俳句の (函数型文法による) 形式的解釈は、

$$H = M (K 1, R, K 2)$$

のように表すことができる。

解釈の逆は、生成である。俳句は 2 つの核文 K 1 と K 2 から、関係 R を加味して形式的に生成す

ることができるとも言える。

つまり俳句の生成は、核文K1とK2にメタ文M()を適用する(=作用させる)ことである。

K1およびK2は、俳句Hと比べると理解しやすい平易な単純文である。したがって、上記の変換は、平易な核文を、風雅に凝縮された俳句に翻訳する過程、と見なすこともできる。この翻訳は、同一言語内翻訳(つまり書き換え)である。

逆に俳句Hから核文K1とK2、そして核間関係Rを抽出することは、俳句Hにメタ文M()の逆関数 $M^{-1}()$ つまり $invM()$ を施すことに相当する。

$$invM : H \rightarrow (K1, R, K2)$$

逆関数 $invM()$ を導出することがすなわち俳句の形式的解釈となる。

俳句の形式的解釈は、実は俳句の翻訳とも見なせる。この場合、翻訳は同一言語内翻訳であっても異言語(たとえば英語)への翻訳であってもよい。要するに限界まで凝縮変形されている擬似文を、通常の文や平易な文に還元することが、すなわち翻訳である。

再論してまとめると、

$$H = M(K1 \ R \ K2)$$

という変換により、自明の意味を持つ核文K1およびK2から、美的に凝縮された俳句Hが翻訳出力される。

核文の異言語への翻訳結果 $Tran(K1)$ や $Tran(K2)$ を経由して、もとの俳句Hの異言語への翻訳 $Tran(H)$ を形式的に入手することも可能である。この場合は、俳句の形式的意味解釈を、異言語により行ったと見なすことができる。

具体例を示す。

1) 古池や蛙飛び込む水の音

松尾芭蕉(春)

人口に膾炙しているなので、取り上げた。古来や

かましく言われるほどの名句ではないと考える専門家も少なくない(伊藤正雄(1976-4) p.28)。筆者の評価は保留するが、俳句の構造を語るときには便利な句であることは間違いない。

(伊藤正雄(1976-4) p.28)はさらに、この古池の句について「芭蕉の俳諧がもはやいはゆる滑稽の域を脱皮して、自然を素直に詠む後世の俳句の性格をはっきり示しているところに歴史的価値があるといへよう」と述べている。つまり名句とは思わぬが、「自然を詠む」という俳句の方向性を確立した句として高く評価すると述べている。

2つの核文K1とK2は：

K1 = 古池がある(見える)

K2 = 蛙が池に飛び込む水音がする(聞こえる)

R = K2という事象の場所(場面)をK1が表現。

上記のRの解釈は常識的で妥当に思われるが、そうではない超現実的な解釈も許容される。俳句は仮想空間を展開する芸術であるから、自由な解釈を必然的に誘起する。

2) 菖蒲茸く古きしきたり今何処

透舟(初夏)

K1 = (初夏には)(屋根に)菖蒲を茸くというしきたりがあった

K2 = 古いしきたりは何処かに行ってしまった。

R = K1をK2が言及(回顧)している。

3) 夏立つや光も風もみどり色

透舟(初夏)

K1 = 夏立つ(夏が来た)

K2 = 光も風も緑色に見える

R = K1と判断する理由をK2が述べる。

2)と3)は、文字通りの拙句であるが、説明しやすいので記載した。

4) 軒を出て犬寒月に照らされる
藤沢周平 [20]

K 1 = 犬が軒を出る
K 2 = 犬が寒月に照らされる
R = K 1 という事象の結果が K 2 である, K 1 と K 2 が時系列で並んでいる.

5) 桐の花踏み葬列が通るなり 同上

K 1 = 葬列が桐の花踏み
K 2 = 葬列が通る
R = K 1 と K 2 が同時に起きる.

6) 死火山の朱の山肌冬日照る 同上

K 1 = 死火山の山肌は朱色
K 2 = 死火山の山肌に冬日が照る
R = K 1 と K 2 が並列して [死火山を] 描写.

7) 風出でて雨後の若葉の照りに照る 同上

K 1 = 風が出る
K 2 = 雨後の若葉が照りに照る
R = K 1 と K 2 が並列して [若葉を] 描写,
K 1 が原因で K 2 が結果と観るも可.

8) 聖書借り来し畑道や春の虹 同上

K 1 = 聖書を借りて来る
K 2 = 畑道を歩く
K 3 = 畑道に春の虹がかかる
R = K 1, K 2, K 3 すべて [聖書と作者をめぐる] 描写, 情景描写から作者の思念が滲み出ている.

Ⅲ 断片文の浅い解析

俳句のような断片文の解析は, 語の列を左から右に走査して, 適当な単位に切断しラベル (標識) を付与するという形式で行う. 句構造のような立体的構造の認識 (抽出) は行わない. 頑健かつ単純な言語処理であるので, パターンマッチング処理, あるいは有限状態マシン (FSM) により実

行できる (Saraki, M. and Nitta, Y. (2005)).

浅い解析処理プログラムの概略は下記のように要約できる. プログラムは現段階で未完であるので, 解析実験は手作業 (思考実験) で行っている.

第1ステップ: 断片文 (俳句) Hを入力する.

第2ステップ: 文節 (内容語+機能語) に切断し, それに品詞ラベルを付与する. この処理には既存のソフトウェア (たとえば茶筌) などの利用を考えている.

第3ステップ: 文節+品詞の列を, 文節+意味的役割子の列に変換する. この変換においては, 品詞列と意味的役割子列の対応規則 (パターン規則) を参照する. 品詞情報には, 語の意味分類コードなどが含まれている.

第4ステップ: 文節+意味的役割子の列を, 2つの核文 K 1 と K 2 に分割する. 核文には意味分類コードを付与する. 核文の個数は, 1個であること (つまり K 1 だけ) もあり得る. また, 核文が3個 (つまり K 1, K 2, K 3) 存在することも稀にある.

第5ステップ: 核文 K 1 (意味分類コード) と核文 K 2 (意味分類コード) から, その関係 R を推論する. 推論には核文関係パターンの推論規則および背景知識を使う.

第6ステップ: K 1 + R + K 2 から, 平文 S を合成 (生成) する. この S が, 当初の俳句 (断片文) H の形式的解釈である.

以下に, 意味的役割子コードの例を, 表 1 として与える. 品詞分類コードおよび意味分類コードの表示は省略するがその部分は, 例題で示す.

表1 rvd 意味的役割子 (Semantic Role Marker) (新田義彦 (2003-10) から引用)

意味的 役割子	コード (略記号)	内容・説明	用例 (該当部分にアンダーライン)
行為者	A Agent	動作の主体, 状態の主体 (無生物主語を含む).	A flat base on which <u>anything</u> rests is called . . . <u>The policeman</u> caught a thief. She has a dream that her son . . .
対象	O Object	属性の主体, 動作の対象, 状態変化の主体.	. . . produces <u>a current of air</u> lose <u>strength</u> make sounds by sending <u>air</u> through
補格・同格	C Complement	記述対象の状態, 様態, Agent, Object の補格または同格.	He is <u>the mayor of our city</u> . Having to do with <u>France</u> as <u>an idiot</u> . . .
関係 [者]・ [関] 与格	R Relating	行為者, 対象と何らかの関係を持つ物, 人, 事など.	According to <u>the radio news</u> , today's weather is partly rainy . . .
道具	I Instrument	手段, 道具, 原因, 条件など.	. . . taking pleasure <u>in giving pain to others</u> . . . destroys <u>with an explosion</u> . . .
始発点	S Source	状態変化の始点, 初期状態.	The spy came <u>from the cold country</u> . They make cheese <u>out of milk</u> .
到達点	G Goal	状態変化の終点, 結果 (受益者を含む).	. . . make <u>angry</u> , . . . go <u>to bed</u> , . . . catch up with <u>them</u> , a piece of ground prepared for <u>plants</u>
場所	L Location	(G, S 以外で) 動作の行われる / 状態の存在する場所.	a small vessel for traveling <u>on water</u> . . . being <u>on the inside</u> . . . stay <u>here</u> for a few seconds
時間	T Time	動作の行われる, または状態の存在する時間・時刻.	We will stay at this hotel <u>for six weeks</u> . I watched TV <u>from ten to twelve o' clock</u> get up <u>at 6 a. m.</u> sharp
目的	P Purpose	動作の目的.	. . . travel by boat for <u>pleasure</u> They stopped <u>to smoke</u> .
状況・事態	X Situation Event	上記以外の状況を示すもの.	a man who is rough and rude <u>in speech and manner</u> We all are <u>in the same situation</u> .
程度・様態	D Degree Mode	動作・状態・状況の程度や様式を示すもの.	This material should be treated <u>with the highest carefulness</u> , because it explodes very easily. Take it <u>as much as possible</u> . You can hit me, if I am wrong, but <u>very gently</u> .
修飾・限定	M Modification Constraint	主体, 対象, 補格, 場所, 時間, など他の深層格に修飾や限定を施すもの. 本来は深層格を持たぬと見なせるが, 処理の便宜を考慮して付与.	A <u>naughty</u> boy was playing with a jackknife. "naughty" は "boy" を修飾している. You must report here tomorrow at seven <u>sharp</u> . "sharp" は "seven" を限定・修飾している. His teacher was an <u>irresistibly attractive</u> lady.
注: 以下は 本研究において暫定的 (試験的) に追加した 意味的役割子 (Semantic Role Marker) である.			
原因・理由	W Wake (わけ)	動作の原因と理由.	Bring my umbrella, <u>as weather is threatening</u> . Please treat me to a lunch; Today <u>I am short of money</u> .
結果・帰結	K Kekka (結果)	動作の結果と帰結.	The robbery shot a clerk <u>to death</u> . He has lived his life in virtuous penury and finally was starved <u>to death</u> . She studied very hard and as a result <u>passed the qualifying examination</u> .

IV 断片文の解析例

1) 雪残る頂(いただき)一つ国境(くにざかい)

正岡子規

(雪残る, Pred) (頂一つ, A) (国境, L)

K 1 = 雪が残る頂が一つある(見える)

K 2 = [見えた場所は] 国境である

R = 事象K 1 発生をした場所がK 2

S = 国境から雪がまだ残る山頂が一つ見えた。

言外の意味, 感懐は形式的意味処理では導出できない。ある種の爽快感, 見晴の良さといった感覚も形式的には導出できない。

2) 些事大事あまさず告げぬ新社員

岩瀬善太

(些事大事, O), (あまさず告げぬ, Pred), (新社員, A)

K 1 = 些事も大事もまさず [上司に] 伝える

K 2 = 新社員である

R = K 2 が K 1 という動作をする / 事件を引き起こす

S = 新社員は些事も大事も漏らさず (取捨選択・重みづけすることなく) [忠実に上司に] 伝える。

新入社員の初々しさ, 四月業務開始時期の会社のオフィスの雰囲気, といった感覚の理解は形式的処理の埒外である。今どきの新社員は業務のエキスパートとしての行動第一であるから, このような初心な伝達行動はとらぬかもしれない。

3) うきわれをさびしがらせよ閑古鳥

松尾芭蕉

(うきわれ, O) (さびしがらせよ, Pred) (閑古鳥, A)

言外に(我, A) (頼む, Pred) (さびしがらせてくれること, O) (閑古鳥, P) が存在する。浅い処理が苦手とする階層性である。

K 1 = 我は憂いている

K 2 = (その我をさびしがらせてみなさい。さびしがらせて憂さを晴らしてください。と郭公に頼む。

R = (K 1 の解決をK 2 で図っている。

S = (郭公よ, 退屈している我にさびしいという感情を引き起こして憂さを晴らしてください。

ひらがなだけの連鎖文は, 浅い解析が苦手とする。

4) 片影の伸びて往診終わりけり

中村四峰

(片影, A) (伸びている, Pred), (往診, A) (終わりけり, Pred)

K 1 = 片影が伸びている

K 2 = 往診が終わった

R = 事象K 1 とK 2 は併存(随伴)している。

S = 道に片影が伸びる頃, ようやく往診が終わった(やれやれ疲れたなあ)。

今は昔の往診ではあるが, 仕事を終えて少しほっとした気持と心地よい疲れがさりげなく詠み込まれている。このような感懐(解釈)も形式処理の埒外である。上記のSは実は形式的解釈ではない。

5) 吹きとばす石はあさまの野分哉

松尾芭蕉

(吹きとばす, Pred), (石, O) (あさまの野分, A)

K 1 = 石を吹き飛ばしている

K 2 = 吹き飛ばしている主は浅間の野分である

R = K 1 への追加説明を K 2 が与える.

S = あさまの野分が石を吹き飛ばしている.

野分, 秋の強風の猛威を浅間の山裾で見た感懐が, 躍動感をもって詠み込まれている. 形式的処理では躍動感抽出できない. 文の叙述構造の抽出は比較的平易な形式処理になじむ.

6) 鴨の背に小さき広さありにけり

今富節子 [21]

(鴨の背, L) (小さき広さ, A) (ある, Pred)

K 1 = 鴨の背を見ている

K 2 = 鴨の背に小さな広さがある

R = K 1 の結果 K 2 に気付いた.

S = 鴨の背には小さな広さがあることに気が付いた. かわいいと思う.

生きとし生きるものの愛おしさを, 小さな背中から観照した句である. 局所を描写して生命現象の極大を語る手腕は見事. そこはかたない哀感の薄い味付けも見事.

7) 割箸の少し抗う秋の風 同上

K 1 = 割箸が [割られまいとして] 少し抗う

K 2 = 秋の風が吹く

R = 事象 K 1 発現しているとき事象 K 2 が発生した

S = 秋の風が吹いている食事時, なぜか割箸が割りにくい.

8) 匂ひなきことのうれしき桜かな 同上

K 1 = 桜には匂いが無い

K 2 = K 1 をうれしく感じる

R = 事象 K 1 の結果 K 2 が発現

S = 桜には匂いが無いが, それを好もしく感じる. 匂いを発散させないのは散り際が潔い桜に相応しい.

9) 一枚をゆらしてあそぶ芋の露 同上

K 1 = 一枚の芋の葉を揺らして遊んでいる

K 2 = 露が芋の葉の上で遊んでいる

R = 事象 K 1 は K 2 の結果である

S = 芋の葉が揺れているが, それは葉の上で露が遊んでいるからだ.

V 断片文解釈の背景知識

俳句のような断片文を, 文法的要素が完備した通常文のように解釈できる理由は, 背景知識の参照にある. 背景知識は歳時記や辞典のようなオントロジーにより与えられる. 成文化しているオントロジー以外に, 句作者がもつ人生経験のような無定形な物語知識も重要である.

形式的解釈の観点から, 個々の句作者の持つ人生物語を集成することは難しいが, その断片を内省的に推量して描写することは可能である. そのごく一部を素描すると下記ようになる.

鴨 ← カモ目カモ科の鳥のうち, 比較的小形の水鳥の総称 [広辞苑 第5版].

← 季語 冬

← 小さい生き物 ← けなげに生きる

← 愛しく思う

割箸 ← 下端から半ばほどの所まで割れ目を入れ, 使うときに割って2本とする箸. 杉・竹などを用いる [広辞苑 第5版].

← 2つに割れぬと困る. 食事の妨げになりじれったく感じる.

桜 ← …春, 白色・淡紅色から濃紅色

の花を開く.八重咲きの品種もある.
古来,花王と称せられ,日本の国花
とし,古くは「花」といえば桜を指
した.・・・[広辞苑 第5版]

- ← 堂々とした花の王様であるが,通常
の花の美質の基本的特徴であるはず
の匂い・香りを周囲に振り撒かない.
- ← それを「散り際の潔さと共に」好も
しく奥ゆかしく感じる.

芋 ← 葉 ← 露をのせている
← ゆれる (俳句的な情景の典型)

VI おわりに

俳句のような日本語断片文は,完備した文法的
属性を具備していないので,伝統的な構文解析処
理にはなじまない.浅い語句解析により,意味的
役割子を付与する処理の方が有効であることを論
じた.そして得られた意味的役割子列から,2つ
の核文K1とK2を抽出すること,2つの核文K
1とK2の意味的関係Rを推定すること,そして
最後にK1+R+K2から,背景的知识を参照し
つつ形式的意味解釈が可能であることを,論じた.

今後は,意味的役割子列から,核文Kiを抽出
同定するためのパターン規則の完備化,核文間の
関係Rの推定に必要なパターン規則の整備,背景
的知識の体系化などを推進したい.

参考文献

- [1] Bentivogli, L. and Pianta, E. (2005) "Exploiting Parallel Texts in the Creation of Multilingual Semantically Annotated Resources: the MultiSemiCor Corpus," *Natural Language Engineering*, 11(3), pp.247-261.
- [2] Hornby, A. S. (1978) *Guide to Patterns and Usage in English*, Oxford University Press (Japanese Translation: Kenzo Ito. (1978). *An English Model, Usage*, Oxford University Publication Office).
- [3] Kinyon, A. (2001) "A Language-Independent Shallow-Parser Compiler," *Proc. 39th ACL Ann. Meeting (European Chapter)*, pp.322-329.
- [4] Lehrberger, J. J. (1978) "Automatic Translation and the Concept of Sublanguage," *Groupe de Recherche en Traduction Automatique (TAUM)*, Universite de Montreal, Canada.
- [5] Lehmann, W. P. (1980) *The METAL System*, Linguistic Research Center, University of Texas, Texas, USA.
- [6] Marcus, M. P. (1980) *A Theory of Syntactic Recognition for Natural Language*, The MIT Press.
- [7] Mihalcea, R. and Simard, M. (2005) "Parallel Texts," *Natural Language Engineering* 11(3), pp.239-246.
- [8] Moon, R. (1987) *The Analysis of Meaning*, in: Sinclair (ed.), Chapter 4, pp.86-103.
- [9] Nitta, Y. et al. (1982) "A Heuristic Approach to English-into-Japanese Machine Translation," in: J. Horecky (ed.), *Proc. COLING 82 (at Prague)* (= *Proceedings of the 9th International Conference on Computational Linguistics*), North Holland Publishing Company, pp.283-288.
- [10] Nitta, Y. et al. (1984) "A Proper Treatment of Syntax and Semantics in Machine Translation," *Proc. of COLING 84 (at Stanford)* (= *Proceedings of the 10th International Conference on Computational Linguistics*), Association for Computational Linguistics, pp.159-166.
- [11] Nitta, Y. (1993) "Referential Structure: A Mechanism for Giving Word-Definitions in Ordinary Lexicons," in: *Language, Information and Computation*, LSK (Linguistic Society of Korea).
- [12] Nitta, Y. (2002a) "A Study of Semantic Typology Patterns and their Transformations," *Economic Review of Nihon University*, 71(4), Nihon University, Tokyo, pp.131-155.
- [13] Nitta, Y. (2002b) "Problems of Machine Translation: From a Viewpoint of Logical Semantics," *Economic Review of Nihon University*, 72(2) Nihon University, Tokyo, pp.23-42.
- [14] Nitta, Y. (2002c) "A Study of Descriptive Language for Sentence Patterns," *Economic Review of Nihon University*, 72(3), Nihon University, Tokyo, pp.35-59.

- [15] Saraki, M. and Nitta, Y. (2005) "The Semantic Classification of Verb Conjunction in the "Shite" Form," *Proceedings of Spring IECEI Conference*, IECEI Japan.
- [16] 新田義彦 (2003-10), 「オントロジー知識を基礎とする質問応答システムの検討」 ("A Study of Question-Answer Systems Based on Ontology Knowledge") 『経済集志』 Vol.73 No. 3, 日本大学経済学部 (2003-10) pp.29-88.
- [17] 新田義彦 (2012-3) 「俳句の意味の形式的解釈の試み」 ("An Essay on a Formal Interpretation of HAIKU"), 電子通信情報学会・2012 総合大会 於 岡山大学, 『A -13 思考と言語 セッション論文集』.
- [18] 新田義彦 (2013-2) 「不言の美文——俳句における省略の機序」 ("Silence of Beautiful Sentences—The Mechanism of Omission in HAIKU"), 電子情報通信学会・『思考と言語研究会 & ことば工学研究会 論文集』 於 明治大学・国際総合研究所.
- [19] 伊藤正雄 (1976-4) 『俳諧七部集 芭蕉連句全解』, 河出書房新社.
- [20] 藤沢周平 (1999-3) 『藤沢周平句集』, 文芸春秋.
- [21] 今富節子 (2005-11) 『多福——今富節子句集』, 角川書店.