

情報理論に依る俳句の分析

藤沢, 偉作
茨城大学工学部

<https://doi.org/10.15017/12730>

出版情報 : 統計科学研究. 2 (2), pp.27-31, 1958-04. Research Association of Statistical Sciences
バージョン :
権利関係 :



情報理論に依る俳句の分析

藤澤 偉 作 (茨城大学・工)

1. 序 論

現在迄俳句に就て論ぜられて来た争柄を情報理論を用いて統一的に導いて来るのが本論文の目的である。

即ち、現在迄経験的に俳句に課せられて来た諸種の条件を情報理論を用いて研究するものであつて、之は亦意味論的情報理論の応用の一つとも信ぜられる。

2. 俳句とは

俳句とは自然、人事よりなるほど時間的に不変な争象の中から美的な争象を抽出して、十七文字により表現しようとする文学と云えよう。

この抽出は美的争象の情報に最大限に伝えるように行われなくてはならない事は明らかである。

抽出された美的争象を標本と呼ぼう。

例えば「古池や、蛙飛込む、水の音」に於ては「古池」「蛙」「飛込む」「水の音」の四つが美的争象よりの標本であつて、之等をこの順で自然界より取り出して来るのが抽出である。

この標本を語に表現した場合に *symbol* と云おう。

「古池」「蛙」「飛込む」「水の音」は日本語の *symbol* である。

「や」は俳句としての美的調和を保つために存在するもので *symbol* とは云わない。

一句の *symbol* 数は 5 ~ 3 で 4 が一番多い。

3. 基本公式

今俳句に使用され得るすべての *symbol* 数を N とする。

いくつかの *symbol* を並べて美的争象を表現するわけであるが、亦一番目に並

び得る *symbol* 数は N が一般には N より小さかるう。

この数を N_1 , 又ある *symbol* が選ばれる確率を p_1 とする。

オ一番目の *symbol* が定まればオ二番目に並び得る *symbol* 数が定まる。

この数を N_2 , ある *symbol* が選ばれる確率を p_2 とする。

以下同様にしてオ三, オ四, ……オ n 番目の *symbol* に対して N_3, N_4, \dots, N_n , 及び p_3, p_4, \dots, p_n が定まる。

「春雨」と云う *symbol* は良く用いられるから p_i が大である。

具体的に p_i がいくらであるかは各 *symbol* に就いて定め難いが, 以下の議論には差支えない。

さて n 個の *symbol* の排列として一つの俳句が形成される確率は $p_1 p_2 \dots p_n$ であって, 情報量は

$$-\log_2 p_1 p_2 \dots p_n = -\sum_{i=1}^n p_i \dots \dots \dots (1)$$

である。

之を基本公式と称する。

俳句を味わう人はその情報とそれにより引き起される自己形成の情報により一つの美的争象を形成するわけであるが, これを鮮明且つ大ならしめるには与えられた情報量が大なる程宜しいわけである。

このためには (1) より n が大なる程, 各 p_i が小なる程良いわけであり, n を大ならしめるにはなるべく文字数の少ない *symbol* を, p_i を小ならしめるには, 人々に今迄余り用いられなかったか, 亦は用いられなかった *symbol* を用いれば良いわけで, さかのぼれば人々に今迄気すかれなかった美的争象の発見, 抽出に帰する。

正岡子規は p_i の大なる俳句を「月並な句」と呼んでいる。

又 n を大にするため「や」などを用いたり, 通例の日本文法を無視した独得の *symbol* の排列法が生れたわけであろうが, 之は基本公式 (1) よりは当然結論される。

4. 季について

季とはある *symbol* の中に四季のどれかの意味が含まれている事である。

今3の N_1, N_2, \dots, N_n を季を考えない場合の *symbol* 数とし, 簡単のために $p_i = \frac{1}{N_i}$ 即ち N_i 個の *symbol* 中より一つの *symbol* がオ i 番目に排列される

確率を同一と見なす。

この場合基本公式により情報量は

$$-\sum_{i=1}^n \log_2 \frac{1}{N_i} = \sum_{i=1}^n \log_2 N_i \quad \dots\dots\dots (2)$$

次に季を考えた場合に、季のある symbol 数と季のない symbol 数との比を i に無関係の $p : q$ ($p + q = 1$) と仮定する。

今 n symbol 中 $i-1$ 番目は季がなく、 i 番目に初めて季のある symbol が並んだとすれば (1) に於て

$$p_t = \frac{1}{qN_t} \quad (t=1, 2, \dots, (i-1)) \quad p_i = \frac{1}{qN_i}$$

又 $i+1$ 番目以下の symbol は無季か i 番目と同季でなければならないから

$$p_s = \frac{1}{qN_s + \frac{p}{4} N_s} \quad (s = i+1, \dots, n)$$

ここで pN_s 個の季のある symbol は四季に 4 等分されると仮定してある。

依つて (1) よりこの場合の情報量は

$$-\log_2 \left(\frac{1}{qN_1 \cdot qN_2 \cdots qN_{i-1}} \cdot \frac{1}{pN_i} \cdot \frac{1}{qN_{i+1} + \frac{p}{4} N_{i+1}} \cdots \frac{1}{qN_n + \frac{p}{4} N_n} \right) =$$

$$= \sum_{i=1}^n \log_2 N_i + \underbrace{(i-1) \times \log_2 q + \log_2 p + (n-i) \log_2 \left(q + \frac{p}{4} \right)} \quad \dots\dots\dots (3)$$

(3)-(2) より両情報量の差は波形の部分となる。

波形の部分は負であるから之が 0 に近い程 (3) は (2) に近いわけである。

波形の部分を一つにまとめると $\log_2 (1-p)^{i-1} p (1-\frac{3}{4}p)^{n-i}$ で、真数が 1 に近い程全体は 0 に近くなり、このためには先ず $i-1=0$ 、即ち $i=1$ が必要である。

この時真数 $f = p(1-\frac{3}{4}p)^{n-1}$ となり最大は $p = \frac{4}{3n}$ の時である。

今 $n=4$ とおけば $p = \frac{1}{3}$, $f = \frac{3^2}{4^3} \therefore \log_2 f = 2 \log 3 - 6$ 。

一方季を入れたために四季の一つがわかるから情報量は $-\log_2 \frac{1}{4} = 2$ だけ増す。

故に季を入れた場合の情報量の最大値は

$$\sum_{i=1}^n \log_2 N_i + 2 \log_2 3 - 6 + 2$$

所が $2 \log_2 3 - 4 < 0$ であるから、季を入れた事により無季の時より情報量は若干減少する。しかし季の入った事により自己形成の美的争象が増大するならば

俳句を味わう人々にとって季のあった方が良いともいえる。

以上を総合するに「俳句に季を入れなくとも俳句として成立するものならば、その方が好ましい。

もし季を入れるならば、季の入った *symbol* は一番目におくべきである。

又季の入った *symbol* 数と全 *symbol* との比は 1:3 が最上である。」

との重要法則が導かれたわけである。

5. 理論と実際の比較

以上俳句は如何にあるべきかを基本公式より導いて来たのであるが、之は俳句の集団を取り扱ったものであって個々の俳句の細かな所に至っては力が及ばない。

又季の問題にしても相当荒っぽい仮定の下に行ったが之も仕方のない事であった。

では4で導いた結果と実際とはどの程度に合うか比較してみよう。

俳句に季を入れなくとも良いとの主張はしばしば耳にするが、目下のところ大勢を制するには至っていない。

だが有名な某俳人が九州地方の炭坑に入坑して句を物しようと努めたが、季にこだわったが故に一句も物にならなかったの例はこの問題に深い示唆を与えている。

亦川柳は季はなくも多大の情報を与えているし、多少 *symbol* 数は多いが和歌がよく美的争象を形成している所より見れば時代の流れと共に4の結論の示す方向に行くものと思われる。

次に季の位置に就ては「高浜虚子他編 俳諧歳時記 春之部」よりの調査結果を掲げる。歳時記ならば良句のみと考え、9570句について、「子規以前」「子規以後」に分けて季の位置の状態を調べた。

大体の句は4 *symbol* よりなっているので季がこの4 *symbol* のどの位置にあるかにより分類したのが下表である。(単位%)

季の位置	1	2	3	4	計
子規以前	39.6	8.5	8.3	43.6	100
子規以後	47.9	10.3	2.7	39.1	100

之によって眺めれば季の位置は時代と共に左に移動していることがわかる。

之は句の数の少い時代にあつては季の位置にかかわらず情報量を大ならしめ得

たが、数が多くなるにつれて情報量の大きなる句を作るには4の理論の示すように季を左に移動せざるを得なくなったものと思われる。

更に極言すれば季が左に移動するにつれて、俳句が古来の法則の下で作られる限り、俳句は滅亡への方向に進んでいると云えよう。

最後に季のある *symbol* 数対全 *symbol* 数の比が 1:3 であるか否かは余りにも調査困難であるために目下は何とも云えないのは残念であるが、その中に何らかの方法で確かめて見たいと思っている。

6. 俳句の将来

基本公式によれば各 p_i が小なる程良いわけであるが、之が小なるためには N_i が大なることが必要且つ充分条件となる。

依って俳句が長らくその生命を維持するためには俳句に用いる *symbol* 数を出来る限り大にすることが必要である。

このためには古来の伝統を墨守せず、絶えず今運用いられなかった言葉を採用し、且つ亦排列法もなるべく自由にすべきであると結論される。

五七五の定型を破ろうとしたり、俗語を取り入れたり、無季を主張したりする運動はこの点より鮮かに説明される。

7. 結 び

以上6項目に亘って基本公式の導入とそれよりの数学的帰結としての俳句の概要を論じたわけであるが、勿論文学的なるものであるから、数学の一つの公式よりそのすべてを結論することは不可能且つ無暴と思われるが、一方俳句の主流問題が一つの基本公式より導かれたと云うことは、情報理論応用の一分野を開くものと自負し、敢てこの一文を掲げたわけである。

多大の御批判を期待しています。

参 考 文 献

- [1] 高浜虚子他：俳諧歳時記 春之部 改造社 (1947)。
- [2] D. A. Bell: *Information Theory*, Sir Isaag Pitman & Sons, Ltd. (1953)
- [3] 中島斌雄：俳句, 朝日新聞社 (1956)