

属性比喩の理解のための計算モデル

内山 将夫 板橋 秀一  
筑波大学

1 はじめに

本稿では、修飾表現の情報理論的モデルを提案し、それに基づいた属性比喩の解釈を述べる。属性比喩の典型は、「狼のような男」という句である[1]。この句は、男を狼に喩えることにより、男の獠牙さを強調している。このように、喩詞(喩える言葉)と被喩詞(喩えられる言葉)との関係のうちでも、特に、両者の意味を構成する属性間の関係により意味が決まるような比喩を属性比喩という。

我々は、文献[2]で、情報理論に基づいた属性比喩の計算モデルを提案した。本稿では、そのモデルを拡張し、修飾表現全体のなかに属性比喩を位置付ける。

2 修飾表現の情報理論による定式化

次の二つの観点から修飾表現を特徴付ける。

- (他の修飾表現に対する) ニュース性、
- (被修飾語に対する) 明瞭性。

まず、名詞の意味を定義し、次に、ニュース性と明瞭性とを定義する。

2.1 名詞の意味

名詞や名詞句の内包的意味[3]を概念と呼び、それを図1のように素性構造で定義する<sup>1</sup>。

〈概念〉 := (〈概念名〉 (〈属性〉 +))  
 〈属性〉 := (〈属性名〉 ((〈属性値〉 〈確率〉) +))

図1: 概念の定義

<sup>1</sup>概念を構成する属性と属性値及びそれに付与する確率の決定法は本稿の考察外である。

ある概念を  $C$  とし<sup>2</sup>、その属性  $F_i$  が属性値  $A_j$  からなるとすると、 $A_j$  の確率  $p(C.F_i.A_j)$  は次の式を満たす。

$$p(C.F_i.A_j) \geq 0,$$

$$\sum_j p(C.F_i.A_j) = 1.$$

図2に概念の例を示す。また、概念  $C$  の属性  $F_i$  のエントロピー  $S(C.F_i)$  は次式である。

$$S(C.F_i) = -\sum_j p(C.F_i.A_j) \lg p(C.F_i.A_j),$$

ただし、 $\lg$  を2が底の対数とする。 $S(C.F_i)$  を概念  $C$  の属性  $F_i$  の曖昧さと呼ぶ。

(狼	((性格	((荒荒しい . 0.9)(大人しい . 0.1)))
	(毛深い	((yes . 0.99)(no . 0.01))))
(男	((性格	((荒荒しい . 0.6)(大人しい . 0.4)))
	(毛深い	((yes . 0.01)(no . 0.99))))
(男の子	((性格	((活発な . 0.7)(大人しい . 0.3)))
	(外見	((若い . 0.99)(老いた . 0.01))))
(女の子	((性格	((活発な . 0.3)(大人しい . 0.7)))
	(外見	((若い . 0.99)(老いた . 0.01))))

図2: 概念の例

一般に、修飾表現と被修飾語とは、外延が同じであり、内包が異なる。これは、同一の確率変数における確率分布の変動、つまり、修飾語を条件とする条件付き確率と考えることができる。すなわち、修飾表現  $NP$  と被修飾語  $N$  について属性  $F_i$  の属性値  $A_j$  の確率は、修飾語を  $AP$  とすると、

$$p(NP.F_i.A_j) = p(N.F_i.A_j|AP)$$

という関係にあると解釈できる。これに基づいてニュース性と明瞭性とを定義する。

<sup>2</sup>混同の恐れがない限り、名詞や名詞句と概念とを同一の記号で表わす。

## 2.2 ニュース性

被修飾語  $N$  について属性  $F_i$  の属性値  $A_j$  の確率が  $p(N.F_i.A_j)$  のとき、修飾表現  $NP$  の属性  $F_i$  のニュース性  $N(NP.F_i)$  を次式で定義する。

$$N(NP.F_i) = -\sum_j p(NP.F_i.A_j) \lg p(N.F_i.A_j) - S(N.F_i)$$

このとき  $p(N.F_i.A_j) \geq \epsilon > 0$  を仮定する。 $\epsilon$  は最小の確率であり、本稿では  $\epsilon = 0.01$  である。

ニュース性は獲得された平均情報量と獲得を期待された情報量(エントロピー)との差である。例えば、被修飾語の属性において確率が  $p(N.F_i.A_j)$  であった属性値が修飾表現において確率1であったとすると、このときに獲得された情報量は  $-\lg p(N.F_i.A_j)$  である。しかし、一般には、修飾表現において属性値が一つには確定しない。そのときには、 $N(NP.F_i)$  の第一項は獲得された平均情報量となる。

これは、修飾表現により限定された外延が、どのくらい稀な事象かを表わしているの、(他の修飾表現に対する)ニュース性と呼ぶ。

## 2.3 明瞭性

被修飾語  $N$  と修飾表現  $NP$  の属性  $F_i$  に関する曖昧さを、それぞれ、 $S(N.F_i)$  と  $S(NP.F_i)$  とする。 $S(NP.F_i)$  は  $N$  に修飾語を付加した結果の曖昧さであるので、 $S(N.F_i)$  と  $S(NP.F_i)$  との差は、修飾表現により与えられた被修飾語の内包に関する情報であり、曖昧さの減少、すなわち明瞭さの増加である。(被修飾語  $N$  に対する)修飾表現  $NP$  の属性  $F_i$  の明瞭性  $C(NP.F_i)$  を次式で定義する。

$$C(NP.F_i) = S(N.F_i) - S(NP.F_i)$$

ニュース性と明瞭性は共に負にもなる。

## 2.4 属性比喩の特徴

属性比喩の特徴は、その内包的意味が喩詞と被喩詞から合成できることである。喩詞と被喩詞とはそれぞれ名詞であるので、その概念を  $V$ ,  $T$  とする。すると属性比喩  $M$  を構成する属性  $F_i$  の属性値  $A_j$  の確率

$P(M.F_i.A_j)$  は次式であると仮定する。

$$p(M.F_i.A_j) = \frac{\alpha}{\alpha + \beta} p(V.F_i.A_j) + \frac{\beta}{\alpha + \beta} p(T.F_i.A_j) \quad (1)$$

$\alpha$  と  $\beta$  は、喩詞と被喩詞の属性値が比喩の属性値に与える影響(重み)を示す。この値は次の三つの要因により決まると仮定する。

### ● 語彙的要因

同一概念でも属性値が異なれば重みが異なる場合がある。

### ● 構文的要因

喩詞には典型的な言葉が使われるため、一般に、 $\alpha \geq \beta$  であろう。

### ● 文脈的要因

比喩が使用される文脈によっては、 $\alpha, \beta$  の値が語彙的要因と構文的要因だけでは決まらない。

本稿では  $\alpha = \beta = 1$  とする。

## 3 ニュース性 / 明瞭性と修飾表現との関係

図3にニュース性 / 明瞭性と修飾表現との関係を示す。図3は、属性値が二つである任意の属性について、被修飾語の属性の確率分布と修飾表現のそれが一定の関係にあるときの、ニュース性(News)と明瞭性(Clarity)の値をプロットしたものである。

図3において、“Minimum”で示されるプロットは、被修飾語と修飾表現とが確率的に独立であるとし、それぞれの属性の確率分布をランダムに選択したときの三万組に対して、ニュース性が最小であったもののプロットである。

また、“Literal Expressions”は、被修飾語の属性値の確率分布をランダムに設定し、一つの属性値を選んで、その属性値の修飾表現における確率を1としたときのニュース性と明瞭性との関係である(五千組)。例えば、“赤い花”という修飾表現は、被修飾語の色に関する属性値のなかで“赤”が修飾表現において確率1となるため、この範疇に入る。このプロットにおいて、被修飾語の属性値のうちで、確率の大きいほうが、修飾表現において確率1ならば、そのときの

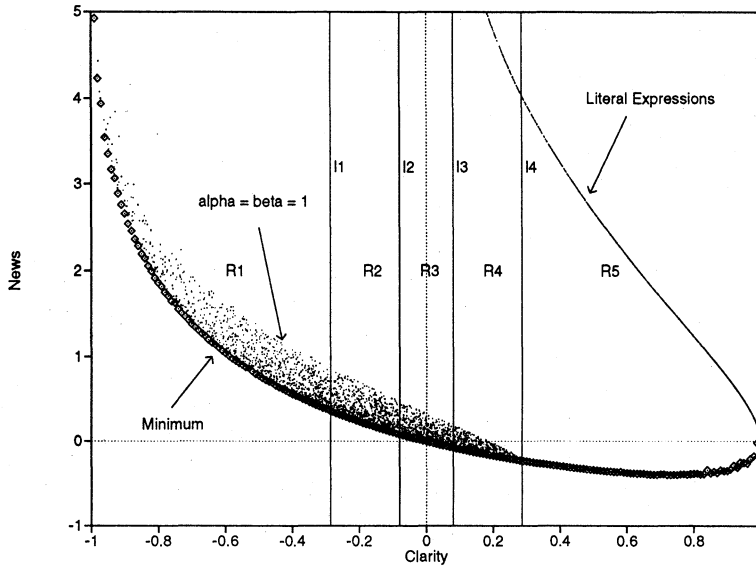


図3: 断定的な表現と曖昧な表現におけるニュース性と明瞭性

ニュース性は最小値になり、“Minimum”の上に落ちる。したがって、“Literal Expressions”が示すプロットは、確率の小さいほうの属性値が修飾表現において確率1であったときのものである。なお、どちらの場合でも、修飾表現の曖昧さは0であるため、明瞭性の値は被修飾語の曖昧さ( $\geq 0$ )に等しい。

“alpha = beta = 1”で示されるプロットは、修飾語と被修飾語の属性値の確率分布が独立であるとした五千組に対して、修飾表現の確率を(1)式において $\alpha = \beta = 1$ として求め、ニュース性と明瞭性とを計算した結果である。

修飾表現の属性値のうちの一つが確率1となり、残りが0となるような修飾表現を断定的な表現と呼び、修飾表現の属性値の確率が(1)式で決まるようなものを混合表現と呼ぶ。すると断定的な表現では、どの属性値が選択されたかが問題なので、ニュース性が重要となる。(どの属性値が確率1となっても、明瞭性の値は変わらない。)それに対して、混合表現では、どの属性値も特には選択されないため、ニュース性は重要ではない。明瞭性が重要となる。

図3は、四本の直線により五つの領域に分割されている。直線I4は明瞭性が0.28の位置にある。被修飾

語において、二つの属性値の確率のうち一つが0.95であり、修飾表現が断定的であるなら、明瞭性は0.29になる。また、混合表現(“alpha = beta = 1”で示されるプロット)における明瞭性の最大値は0.27である。したがって、I4は断定的な表現と混合表現とを区別する。I4の右側、領域R5に位置する修飾表現が断定的な表現である。

直線I3は、明瞭性が0.08の位置にある。これは、被修飾語において、二つの属性値のうちの一つの確率が0.99であった場合の断定的な表現における明瞭性である。また、直線I2は縦軸に関してI3と対称である。この二本の直線により区切られる領域R3に位置する表現は、修飾表現として不適当であると考えられる。例えば、図2の概念によると、“若い男の子”や“老いた男の子”がそうである。前者は、修飾するほどの価値がなく(ニュース性が低く)、後者は表現の対象が考えにくい(ニュース性が高すぎる)。つまり、R3には、当り前のことか有りえないことを述べた表現が位置する。

直線I1は、縦軸に関してI4と対称である。対称である必然性はないが、これにより分けられる領域R1には、明瞭性の値が低いため修飾表現として不適当な

表現が位置すると考えられる。

残りの領域  $R2, R4$  が属性比喩の領域である。ここに位置する表現の明瞭性は適度なものである。それに加えて喩詞と被喩詞とが存在するという言語形式から、当該の表現が属性比喩として適当な表現であると判断される<sup>3</sup>。

#### 4 属性比喩と明瞭性

前節で述べたように、属性比喩においては明瞭性が重要となる。表1に図2から構成される表現を明瞭性と共に示す。表現の属性値の確率は(1)式において  $\alpha = \beta = 1$  として求めた。

領域	表現, 属性	明瞭性
R4	狼のような男, 性格	+0.16
R3	男の子のような女の子, 外見	0
	女の子のような男の子, 外見	0
R2	男の子のような女の子, 性格	-0.12
	女の子のような男の子, 性格	-0.12
R1	男のような狼, 性格	-0.32
	男のような狼, 毛深い	-0.92
	狼のような男, 毛深い	-0.92

表1: 属性比喩と明瞭性

属性比喩の特徴のうち、次の二つが表1から読み取れる。

**選択的強調** 被喩詞の属性のうちの幾つかが比喩において強調される。強調される属性は、領域  $R2$  と  $R4$  に位置するものである。例えば、「狼のような男」では「性格」という属性が強調される。「毛深い」という属性は、「狼」と「男」とでは確率分布が大幅に違うため、領域  $R1$  に位置することになり、強調される属性とはならない。

**意味の非対称性** 「 $A$ のような $B$ 」と「 $B$ のような $A$ 」とでは意味が異なる。これには明瞭性の非対称性と

<sup>3</sup>「いかつい男のメガネの色」というような連体修飾表現もここに位置するであろう。この場合には、メガネの色は黒である確率が高くなると思う。

確率分布の非対称性とが関わる。「狼のような男」とは言うが「男のような狼」と言わないのは、前者においては、領域  $R4$  に「性格」という属性が位置するのに対して、後者では、全ての属性が  $R1$  に位置するためである。これは明瞭性の非対称性である。

「男の子のような女の子」と「女の子のような男の子」とでは、「性格」という属性が領域  $R2$  に位置する。これらの明瞭性に差はない。しかし、(1)式で計算される属性値の確率を見ると、「男の子のような女の子」では、「性格」という属性のうち「活発な」という属性値の確率が0.3から0.5に上っている。それに対して、「女の子のような男の子」では、「大人しい」という属性値の確率が上っている。これは確率分布の非対称性である。

#### 5 おわりに

修飾表現の意味は、修飾語を条件とする、被修飾語の意味の条件付き確率であると仮定した。それに基づき、(他の修飾表現に対する)ニュース性と(被修飾語に対する)明瞭性とを定義した。次に、属性比喩の意味が喩詞と被喩詞の意味から合成できると仮定した。そして、属性比喩では、ニュース性と明瞭性の観点のうち、明瞭性が重要であることを示し、属性比喩の特徴のうちで、属性の選択的強調と意味の非対称性とを明瞭性により説明した。

本モデルの妥当性を実証することが今後の課題である。

#### 参考文献

- [1] 諏訪正樹、岩山真：「比喩の計算モデル」, 情報処理 Vol.34, No.5 (1993).
- [2] 内山将夫、板橋秀一：「視点を考慮した比喩の理解」, 情処研報 95-NL-105-8, (1995)
- [3] 白井賢一郎：「形式意味論入門一言語・論理・認知一の世界」, 東京：産業図書, (1985).