

日本語の多義的な並列名詞句の解析

山中信彦 (埼玉大学教養学部)

0 要旨

日本語の統語的に多義的な並列名詞句を人間がどのようにして一義的に解析しているのかに関する認知的なモデルを提案する。モデルは少数の一般的な原則と、それらが適用される順序からなる。新聞記事に含まれる多数の実例によってモデルの有効性を検証する。

1 対象とする並列名詞句の形式 (田村・田中(1987)で対象とされたものと同じ)

「 A_1 の… A_m の X と Y の B_1 の… B_n 」

X 、 Y 、各 A_i 、 B_j ($1 \leq i \leq m$, $1 \leq j \leq n$) は名詞 (接頭辞、接尾辞のついた名詞や複合名詞などを含む) を表す。次のようなものは対象としない。

きのうの答弁と安保国会当時の答弁が食違う (A035051) 相手格の「と」

[[[フィリピンとインドネシア] 両国] 間] の問題だ (A059582)

[スタンドオフのキック] と [スクラムサイドからの突進] を (A001337)

ハノイ市とハイフォン市の軍民と人民軍空軍部隊は (A088328) 並列詞が複数

七回に二塁打した田中と柳田の長打が (A037774) 形式に合わない連体修飾句

大幅賃上げと独自の要求を (K078266) 統語的に名詞と言いがたいもの

2 認知モデル

2-1 前提とする考え方

基本的には言語の線状性に基づき入力される順序に従って部分的に解析を行いながら次を予測し、更なる入力を解析結果に結びつけていくと考える。

然るべき理由がある限り訂正 (バックトラッキング) の余地は残されているが、訂正は少ないほど望ましい。

2-2 「 A_1 の A_2 の… A_m 」の解析

山中(1988:78)の仮説に基づいて行う。

ホテルの玄関の扉→' [ホテルの玄関] の扉'

年賀状の蛇の絵→' 年賀状の [蛇の絵] '

2-3 一般原則と、それらが適用される順序

ア 意味的対応要素に関する原則

並列される名詞同士は意味範疇が近いということは、石綿(1965)、(1968)などで指摘されている。ここでは石綿(1968:155)の考え方に基づき、より一般化した原則をたてる。

ア-1 X と対応すべき、意味範疇の近いものを入力順に ($Y \rightarrow B_1 \rightarrow B_2 \rightarrow \dots B_n$) 探していく。「 Y の B_1 」、「 Y の B_1 の B_2 」などとの対応も調べる。

(1) パパとママの前で (A003677)

(2) 板橋区成増と隣の同区上赤塚で (J137717)

(3) 国民と海外の同志たちに (A041421)

ア-2 X と対応しない Y 、 X と対応するかどうかの判断に迷う Y についてはスタック (A_1 , A_2 , ..., A_m , A_1 の A_2 , A_1 の A_2 の… A_m , A_2 の A_3 , ……) の中から Y と対応すると確信できるものを探す。

- (4) 都バスの後部と都電の前部を (A024860)
 (5) 向こう三か月の予報と寒候期予報を (C154652)

名詞の意味範疇の類似度には色々な段階があり、人間はある値以上の類似度を持つ名詞同士を並列の対応要素と判断できるものとする。特定の文脈がなくても基準値以上の類似度を持つ名詞の組としてはここでは次のようなものを考えている。

同位語： 毛、木綿、レーヨン、テトロン、ポリエステル； 休憩室、勤務室
 同義語： マージン、手数料； ショック、驚き； 協会、連盟
 対立語： 量、質； 日中、夜間； 海、空； 母、子

次の原則が必要かどうか検討中である（今回の検証では使わなかった）。

ア－3 ア－1においてXと対応すると確信できるものがある程度探しても見つからなければ、前に戻ってこれまでに出てきたものの中で相対的に意味範疇の近いものを結果的にXの対応要素とみなすことができる（ア－2におけるYについても同様）。

- (6) 企業の中の技術革新の進展と職種の変化とも (C105794)

イ 「束ね」に関する原則

水谷・田中(1972:22)は「解の存在一意性について」などの並列の用法を「並べ」、
 「Recursion theory ト整数論の共演」などの並列の用法を「束ね」と呼んで区別している。
 「XとYのB₁」について言えば、「束ね」の用法では必ず「[XとY]のB₁」と解析されるが、このことはほとんどの場合B₁の性質によって予測できる。一般的に言えば：

B_nがそれ自体のほかに複数の要素の存在を前提とする概念を表している（長尾ほか(1983:3)の「並列句を前方に要求しやすい語」にあたる）ならば、XとB_{n-1}とが並列要素のヘッドとなるように（「[A₁の…A_mのX]と[YのB₁の…B_{n-1}]」のB_n」など）解析する。

- (7) 美と信仰の間に (J177464)
 (8) 政治と国民の直結で (C136874)
 (9) 小林と藤下の二人が (A050296) 「の」がいわゆる同格を表す場合
 (10) 主婦とソロバンの両刀を (A070685) 同上
 少数ながら、上記の原則では予測できない「束ね」の例がある：
 (11) 水と油の感じ。 (K061147)
 (12) 訂正とおわびの文書を (A076715)
 (13) 軽食と喫茶の店が (J178920)

ウ 語結合に関する原則

ウ－1 対応関係が認められる要素が検出された時にはその時点で前後の修飾関係にある要素との語結合を調べる（XとYが対応している場合、「A_mのY」、「XのB₁」など）。その結果、意味論的・語用論的・慣用的に不適格になる解析は除外し、それ以外の解析を候補として残す。語結合が適格であれば、隣接する要素は並列名詞句全体と修飾関係を持つとみなす。

- (14) 政治家の現実主義と経験主義に (C139805)
 (15) ハウス主人の川村さんと非行少年たちが (A086662)
 (16) なにがしかの灰色の過去と今日の生活を (J187366)

Wu and Furugori(1998:10)は英語の“n1 and n2 n3”という構造に関し、n1とn2, n1とn3それぞれの意味的類似性を計算して類似度の高い組を並列対応要素とみなすというやり方をとっているが、ここでは、意味範疇の観点から人間が確信を持って判断できる組が検出されない場合、語結合の観点から判断すると考える。

ウー2 「XのB₁」、「YのB₁」が語結合の点で適格で、かつ「の」が同じ意味関係を表している時「[XとY]のB₁」と解析できる。「A_mのX」、「A_mのY」についても同様の条件のもとで「A_mの[XとY]」と解析できる。

(17) 若さと文部政務次官の肩書きに (C188881)

(18) 国連記念切手と東京支部の扱い# (A053891) 両方が同時に対象になる状況は考えにくい。「同じ意味関係」条件を課せば、「[XとY]のB₁」は排除される。

(19) 人間の幸福と文明の関係を (A043131)

今回の検証ではウー2は適用可能となり次第適用することにしたが、(特に「の」が続く時に)適用を少し保留した方が無駄な解析をしなくてすむような場合がいくつかあった。

(20) 南ベトナムの将軍たちと仏教教会の指導者六人の間に (A074475)

すぐ適用: 「[南ベトナムの[将軍たちと仏教教会]]の…」

少し保留: 「[南ベトナムの[[将軍たち]と[仏教教会の指導者六人]]]の…」

原則相互間の優劣関係: イ>アー1、アー2>ウー2

イは意味範疇の対応の有無に拘わらず適用される。ウー2はアによる訂正を受ける。

原則の適用される順序:

「XとYのB₁」型: アー1→イ→「はい」

→「いいえ」→ウー1またはウー2

「A_mのXとY」型: アー1→「はい」→ウー1

→「いいえ」→アー2→「はい」→アー1

→「いいえ」→ウー2

3 実例による検証

田中ほか(1979)の234~245頁中の、1節で述べた形式に合う名詞句全て(219例)について検証した。名詞の類似度が基準値以上かどうかの判定、語結合が適格かどうかの判定は筆者の直観による。次の結果が得られた。

解析に最も貢献したと考えられる原則(1例に1つ)	アー1	イ	ウー2	計
並列名詞句の形式				
XとYのB ₁	35(1)	38(2)	39(1)	112(4)
XとYのB ₁ のB ₂	3(1)	11	4	18(1)
A ₁ のXとY	26	0	26(4)	52(4)
A ₁ のXとYのB ₁	13	2	12(1)	27(1)
A ₁ のXとYのB ₁ のB ₂	1	0	2	3
A ₁ のA ₂ のXとY	1	0	1	2
A ₁ のA ₂ のXとYのB ₁	3	0	1	4
A ₁ のA ₂ のA ₃ のXとYのB ₁	1	0	0	1
計	83(2)	51(2)	85(6)	219(10)

()内は問題がある例の内数。

問題となる例

(21) 奥さんと奥さんの両親の四人暮らし。(A051487) ' [[奥さんと奥さん]の両親]の…' は誤りなので一見反例だが、ア－1に続いてウ－1を適用するところで語用論的な不適格性が発生するとみなせば必ずしも反例ではない。XとYが同一の場合はB₁が「束ね」を要求する名詞の時に限って(「目と目の間隔」)、' [XとY]のB₁' と解析する。

(22) 力と技術の差だけだったろうか (C155458) イの予測に反し、「束ね」ではないが、ここでの枠組みではやはり' [力と技術]の差' と解析される。

(23) 真実と正義の闘いは (K026424) 同上。

(24) 二つの死球とバントで (A083396) ウ－2の反例。数詞は例外とすべきか。

(25) こんごの運営と都民の協力。(C113835) 同上。「こんご」は「運営」だけに係りそう。

(26) アフリカの自由と統一が (K026424)

(27) ボン政界の危機と混乱は (C161021)

(28) わが国の陶芸発展と海外紹介に (K057799)

(29) 東急と土地の所有者が (K011194)

(30) わざと力の妙技を (C144041)

(31) 期待と不安の日を (A040620)

4 今後の課題

既に触れた、原則の数と適用時期に関する問題のほか、解析がどの時点で確定するかという問題がある。また、本稿のモデルは言語心理学的実験によっても検証する必要がある。

文献

石綿敏雄(1965)「並立助詞『や』『と』の機能」『計量国語学』第32号。

石綿敏雄(1968)「言語の意味と言語情報処理」『電子計算機による国語研究』所収、国立国語研究所。

田中穂積・荻野綱男・荻野孝野(1979)『日本語全品詞列集成 左順篇』電子技術総合研究所推論機構研究室。

田村直良・田中穂積(1987)「意味解析に基づく並列名詞句の構造解析」『情報処理学会自然言語処理研究会資料』59-2。

長尾真・辻井潤一・田中伸佳・石川雅彦(1983)「科学技術論文における並列句とその解析」『情報処理学会自然言語処理研究会資料』36-4。

水谷静夫・田中幸子(1972)「語の並列結合子」『計量国語学』第63号。

山中信彦(1988)「日本語の多義的な名詞修飾構造の解析」『言語研究』第94号。

Wu, H. and Furugori, T. (1998) "A Hybrid Approach for Resolving Ambiguities in Coordinate Structures" 『自然言語処理』Vol. 5, No. 4.

謝辞： 田中ほか(1979)の利用に関して便宜を図って下さった東京工業大学の田中穂積教授に厚くお礼申し上げます。