

日本語文章からの知識の抽出

——保田與重郎における「は」と「が」との機能分析を中心に——

谷口敏夫

はじめに

昨今、コンピュータ上で扱える日本語のテキストが人々の目に触れ、手にすることができるようになってきた。それに付随して、様々な分野で大規模なテキスト批判が生まれている。一方、図書館や図書館学の分野においても、こうした電子化されたテキストに対する基礎的な考究が必要とされてきている。

図書館学において、ハイパーテキストの研究と応用は、⁽¹⁾一般的な市販ソフトウェアによって比較的速やかに進むであろうが、日本語の文章全体を知的に処理するのは、なかなか困難な問題である。これらは、長らく情報工学の分野において、自然言語処理や人工知能として扱われてきた。それは言語情報処理（言語工学）として非常に成果があがっている。

しかし、実際の図書館では、それらの先端科学の果実がなかなか身近なものとはなっていない。それは、言語工学と日本語文法という非常に専門的な領域の二つに交わる所から得られるものであり、従来の図書館学の考え方や方法では、取り組みにくい面が多々あった、ということである。

本稿では、そうした先端科学の一端を、理解できる範囲で図書館学に応用することを目的の一つとしている。このことによって、抄録、索引、知的ハイパーテキスト、知的データベースなどがより深化するものと期待される。また、長い目で見れば、自動分類や自動参考業務などにも、ある程度確かな方法論をもって、適用させることができるのではなかろうか。

さて、本稿では「は」や「が」の日本語における働きを比較的長い文章において調査し、そこからなんらかの知識を簡便に抽出する手続きのための基礎研究を扱った。題材としては、保田與重郎『日本の文學史』の序説⁽²⁾（以下「保田序説」と略する）を選んだ。

そのために、第1章では、対象とした保田與重郎に関しての説明と、文章に対する言語工学的な手法をとりあげ、概略を示した。さらに、日本語の文法全般に対する本稿の立場を概説した。

次に第2章では、基礎研究を円滑に進めるために開発した道具（ツール）の仕組みを解説した。このツールは、本稿のために用意したもので、「は」と「が」⁽³⁾によって、文の構造を視覚的に表現するためのものである。これは「はが文構造表示プログラム」と仮に名付けておいた。また別に、本文を対象とした全文検索のための、島探索プログラム⁽⁴⁾を使用した⁽⁴⁾が、これはすでに発表している⁽⁴⁾ので割愛した。

第3章では、これらのツールを用いて、典型的な「名詞＋は」、「名詞＋が」からの知識の抽出を手技で行い、ハイパーテキスト構築の支援となる可能性を論じた。

第4章では、結果の考察と将来の展望をまとめとした。

いずれも、本稿における方法論を詳細に明示することによって、言語工学の図書館学への応用、取り込みという困難な作業のための一助としたい。

第1章 文章からの知識抽出

1. 1. 保田與重郎

本稿で対象とした文章は保田與重郎（明治43～昭和56：東京帝国大学文学部美学美術史科卒）の戦後の作品『日本の文學史』の序説である。彼は戦前戦後を通して、日本の精神史を文学の立場から一貫して語り継いだ批評家である。その文学史観は、古事記における倭建命、万葉集における大伴家持、新古今和歌集における後鳥羽院、俳諧における芭蕉の、脈々とした流れに日本の精神史を見るものであり、一般的な文学史とは趣をことにする。その文体は格調がたかく、

現代における日本文学の一つの特殊な典型を示しているものである。これを現代の言語工学の、初歩的な手法でどのように扱えるのか、という答えの端緒を見いだすことも本稿の目的の中には含まれている。

1. 2. 言語工学の成果

文章の分析といえば、安本美典は『文章心理学入門』⁽⁵⁾において、すでに昭和40年以前に因子分析法を用いて、森鷗外ら8人の作家の3次元空間上における位置づけを行っている。ここには実験の詳細が分かりやすく記述されている。しかし、本稿は文章そのものをときほぐすことによって、そこから「わかりやすい知識」を抽出することに重きをおくので、この方法論はとらない。統計という客観的な答を得るまえに、一つの作品ないし作家を文章のいろいろな角度からみることにより、文章そのものから知識を導き出すための方法論を確立することに目的をもつ。

文章と知識に関して、言語工学という言葉がある。これは『言語工学』⁽⁶⁾で使われており、同書は昭和58年にまとめられた。この「第2章 言語の統計的性質」「第3章 文字・単語列の処理」「第7章 意味に関する理論」を順番にひもとき、それとは別に『画像と言語の認識工学』⁽⁷⁾での「6. 単文の構文・意味解析」「7. 意味・文脈情報を用いた文の解析」などで、言語工学の具体的な手法を知ることができる。1992年には、「専門用語辞典の自動的ハイパーテキスト化の方法」⁽⁸⁾によって知識そのものを抽出する実験がおこなわれた。その具体的な方法論は、「ハイパーライブラリー」⁽¹⁾の中ですでに述べた。

『言語工学』の著者は世界における人工知能、自然言語処理、機械翻訳の代表者の一人であり、この分野における業績は多岐にわたる。その成果をすべて本稿に導入することはできない。しかし、最近の事例翻訳（類推による翻訳）⁽⁹⁾における考え方の多くの成果は、文ないし文の一部（句）を一つの知識の塊としてあつかい、膨大な知識同士のパターンマッチングをはかる、新しい方法論を導き出したものであり、本稿の方法論の仮説部分がすでに言語工学では手を付けられているという期待がある。

1. 3. 日本語文法学の成果

日本語の文法を言語工学に生かすには、文法を専門的に研究しておられる人達の成果をよく学び研究しなければならない。しかし、日本語の文法研究といっても様々な側面があって、その違いを分野外から正確に把握することさえおぼつかない。

そうした中で、テキストを先述の言語工学の中で扱う場合、何らかの指標を自他ともに定義つけておかないと、本稿自体の混乱をまねく。よってここでは、言語工学的にテキストを解釈していく際に、もっともわかりやすい方法論であったという経験的な事実から、次のような観点の指標を打ち出しておく。

①基準とする文法書

益岡隆志、田窪行則『基礎日本語文法 一改訂版⁽¹⁰⁾』。これは語彙索引と事項索引とが整備され、現代日本語の文法用語の定義がわかりやすく迅速に理解できる。益岡隆志には、すでに『命題の文法⁽¹¹⁾』『モダリティの文法⁽¹²⁾』があり、両著では文法を包括的にかつ記述的にまとめている。

②文章論として捉える考え方

時枝誠記『日本文法：口語篇⁽¹³⁾』、永野賢『文章論総説⁽¹⁴⁾』。この二つは一般的に後者が前者を継承している関係にある。文法を文章全体の中で捉える試みの一つの流れである。

③言葉の歴史的な考え方

大野晋『係り結びの研究⁽¹⁵⁾』、『日本語の文法を考える⁽¹⁶⁾』。この二つでは、「は」と「が」とを、歴史的な係り結びの観点からわかりやすく説いてあり、混乱が生じたときに指標が得られる。

④言葉を扱う態度

小説『やちまた⁽¹⁷⁾』に表現された本居春庭の文法解析の歴史的事実や、わが国における世界的な業績としての西夏文字解読⁽¹⁸⁾の過程がある。この2例をあげたのは、そこに言葉全体にたいするいとおしみ、緻密な帰納的方法の帰結として、日本語の文法や、未解読文字であった西夏文字が解きあかされていた事実をみたからである。本稿の対象とした保田與重郎もまた、現代日本語においては、

非常に難解な文章に位置づけられているものであり、公理的、演繹的な方法論は、望みえないものと推量している。よって、ことばとしての文学のやちまたに入り込む要素は非常にたかい。その際、例にあげた未知のものに対する態度と方法論は一定の指針を示すものと考えた。

なお、②と④に関する本稿での基本的な考え方は「2001年、フルテキストデータベースの未来⁽¹⁹⁾」ですでに述べた。また「は」とか「が」のような平仮名1文字に着目した直接の経緯は、「源氏物語」の写本間での、文字一つの異同による、文意がねじれるほどの事実を学んだことによる。

1. 4. 知識の定義と知識の提示

本稿では、知識とは文章中から

「(比較的意味の明瞭な) 名詞＋{は | が}＋述部」という形の文を抽出したものに限定する。その提示例を次の図1に列挙しておく。詳細は3章で述べた。

大伴氏は	最も斬新な教養の中でくらしてみた
親房卿が	「神皇正統記」を著された
保田與重郎は	鹿持雅澄翁の「万葉集古義」に学んだ
保田與重郎は	伝教大師を知つてみた
保田與重郎は	戦後水田を復元して、そこへ米を植ゑた

図1 知識の提示

第2章 日本語文章の処理方法

ここでは、文章から「は」と「が」を抽出し、それを用いて文の構造を視覚的にわかりやすく表示するプログラムについて説明する。

2. 1. はが文構造表示プログラム

文章は1文単位でまとめてコンピュータに読み込み、その中で「は」ないし「が」が一つでも出現する文に関して、それぞれを「は判定部」、「が判定部」におくり分析した。各々の判定部では解析結果の文の該当個所にプログラムで特殊なマーク付けをし、最後に「採取部」において、文の構造的な表示をおこなった。まったく「は」、「が」をふくまない文は、「棄却部」において後処理

をさせた。なお文の判定は「。」とし、段落の判定は改行マークとした。

処理中と処理結果の文につけた特殊マークは表1に、また各規則の中で使用した命令等の説明は表2にそれぞれまとめておいた。

表1 文中の特殊マークの種類

「は」に関する、特殊マーク。	
☆	助詞「は」。採取部で最後に付加する。
○	注意を要する助詞「は」。は判定部「規則は(4)」で使用。「保田序説」では、使用例がなかった。
◇	非取扱とした「は」。
「が」に関する、特殊マーク。	
★	格助詞「が」。採取部で最後に付加する。
■	接続助詞「が」。
◆	非取扱とした「が」。

表2 命令と記号

文字列処理その他の命令、記号	使い方の説明
ReplAll (検索文字, 置換文字, 対象文); Repl (検索文字, 置換文字, 対象文); StrPos (検索文字, 対象文); IF 条件 THEN 作業; {は が}	対象文の総ての検索文字を置換する 対象文の検索文字を1度だけ置換する 対象文における検索文字の位置を得る 条件に適った時だけ作業をする は「或いは(OR)」の意味

2. 1. 1. は判定部

ここでの判定によるマーク付けは、規則は(4)を除き、後の採取部で取り上げられない「は」を意味する。すなわち、求める助詞「は」ではないというマーク付けである。助詞「は」を自動判定する一般的な規則は、菅沼明⁽²¹⁾を参考にし、それを規則は(1)～規則は(4)にまとめた。保田序説からは別途規則は(5)を設けた。規則の適用は規則番号の順におこなった。

助詞「は」の区分については、菅沼らは「とりたて詞←題目(提題)+対

照」という構造で捉えている。一方、益岡隆志⁽¹⁰⁾では「は」を{提題助詞, 取り立て助詞}というように並記し, その中での取り立て助詞「は」は主に対比の意味で用いている。これは菅沼での「対照」に照応する。ここでは益岡に従うが, 混乱を避けるために, 大きく助詞「は」の判定という観点から, 稿を進める。

規則は (1)

文中の「をは」における「は」は, 助詞「は」ではないと考え, 文中の総ての「をは」にマーク付けをする。

```
ReplAll('をは', 'を◇', 文);
```

規則は (2)

文の先頭が「は」なら, 助詞「は」ではないと考え, 先頭の「は」だけをマーク付けする。

```
IF StrPos('は', 文) = 先頭 THEN Repl('は', '◇', 文);
```

規則は (3)

文中の「は」の直後が「ん」撥音, ないし「つ」促音ならば, 助詞「は」ではないと考え, 文中の総ての「はん」「はつ」にマーク付けをする。

```
ReplAll('はん', '◇ん', 文); ReplAll('はつ', '◇つ', 文);
```

規則は (4)

「はじめ (始め)」や「はじまり (始まり)」とまぎれやすい助詞「は」に続く用語採用のためのマーク付けである。この規則は, 規則は (5) との関係が深く, 結果に注意を要するので, マーク○を使いそのまま出力結果に反映させた。ここでは, 「はじめつ」を「は自滅」の意味で使用しているが, 文章としては「一方始めつつも, 他方～」というような文も想定できる。ただし保田序説では, 規則は (4) の該当例は一例もなかった。

```
ReplAll('はじめい', '○じめい', 文);      ~は自明
ReplAll('はじめじめ', '○じめじめ', 文);  ~はジメジメ
ReplAll('はじめつ', '○じめつ', 文);      ~は自滅
ReplAll('○じめつつ', '◇じめつつ', 文);   (非助詞「は」のマークを付けた)
```

ReplAll(‘はじめん’, ‘○じめん’, 文);	～は地面
ReplAll(‘はじめい’, ‘○じめい’, 文);	～は地米
ReplAll(‘はじめえ’, ‘○じめえ’, 文);	～は自前
ReplAll(‘はじめく’, ‘○じめく’, 文);	～は字幕
ReplAll(‘はじめま’, ‘○じめま’, 文);	～はじめま (我侷の意味)
ReplAll(‘はじめわり’, ‘○じめわり’, 文);	～は地回り
ReplAll(‘はじめん’, ‘○じめん’, 文);	～は自慢

規則は (5)

保田序説で採取した助詞「は」ではない「は」。これに該当する文中の総ての「は」に、マーク付けをする。以下の用語には、歴史的仮名遣いと、保田自身の文体の特徴が表れている。一般的に保田の文章は初期においては、非常に難解な漢語が多いが、戦後の『現代崎人傳』以降は、⁽²²⁾ひらかな遣いが多く見られる。以下のように59例の規則を設けた。

ReplAll(‘あはせ’, ‘あ◇せ’, 文);	ReplAll(‘はかり’, ‘◇かり’, 文);
ReplAll(‘あはな’, ‘あ◇な’, 文);	(漢字事例: 「計り知れない」)
ReplAll(‘あはぬ’, ‘あ◇ぬ’, 文);	ReplAll(‘はげし’, ‘◇げし’, 文);
ReplAll(‘あはれ’, ‘あ◇れ’, 文);	ReplAll(‘はこぶ’, ‘◇こぶ’, 文);
ReplAll(‘あらは’, ‘あら◇’, 文);	ReplAll(‘はじめ’, ‘◇じめ’, 文);
ReplAll(‘あるひは’, ‘あるひ◇’, 文);	ReplAll(‘はずみ’, ‘◇ずみ’, 文);
ReplAll(‘いはない’, ‘い◇ない’, 文);	ReplAll(‘はたから’, ‘◇たから’, 文);
ReplAll(‘いはねば’, ‘い◇ねば’, 文);	ReplAll(‘はたらき’, ‘◇たらき’, 文);
ReplAll(‘いはれ’, ‘い◇れ’, 文);	ReplAll(‘はて’, ‘◇て’, 文);
ReplAll(‘かかはら’, ‘かか◇ら’, 文);	ReplAll(‘はらばひ’, ‘◇らばひ’, 文);
ReplAll(‘かぐはし’, ‘かぐ◇し’, 文);	ReplAll(‘はるか’, ‘◇るか’, 文);
ReplAll(‘かなはな’, ‘かな◇な’, 文);	ReplAll(‘はれた’, ‘◇れた’, 文);
ReplAll(‘かなはぬ’, ‘かな◇ぬ’, 文);	ReplAll(‘はれる’, ‘◇れる’, 文);
ReplAll(‘かはり’, ‘か◇り’, 文);	ReplAll(‘はやり’, ‘◇やり’, 文);
ReplAll(‘がはり’, ‘が◇り’, 文);	ReplAll(‘ふさはし’, ‘ふさ◇し’, 文);
ReplAll(‘かまひは’, ‘かまひ◇’, 文);	ReplAll(‘まぎは’, ‘まぎ◇’, 文);
ReplAll(‘こだは’, ‘こだ◇’, 文);	ReplAll(‘もはや’, ‘も◇や’, 文);
ReplAll(‘そぐは’, ‘そぐ◇’, 文);	ReplAll(‘わけはな’, ‘わけ◇な’, 文);
ReplAll(‘はかない’, ‘◇かない’, 文);	(「訳はない」の慣用扱いとした)

ReplAll(‘わらはれ’, ‘わら◇れ’, 文);
 ReplAll(‘扱はれ’, ‘扱◇れ’, 文);
 ReplAll(‘現はし’, ‘現◇し’, 文);
 ReplAll(‘現はす’, ‘現◇す’, 文);
 ReplAll(‘現はせ’, ‘現◇せ’, 文);
 ReplAll(‘現はれ’, ‘現◇れ’, 文);
 ReplAll(‘淡あは’, ‘淡あ◇’, 文);
 ReplAll(‘味はれ’, ‘味◇れ’, 文);
 ReplAll(‘或ひは’, ‘或ひ◇’, 文);
 ReplAll(‘云は’, ‘云◇’, 文);
 ReplAll(‘言はれ’, ‘言◇れ’, 文);
 ReplAll(‘歌はれ’, ‘歌◇れ’, 文);
 ReplAll(‘行はれ’, ‘行◇れ’, 文);
 ReplAll(‘思はせ’, ‘思◇せ’, 文);
 ReplAll(‘思はな’, ‘思◇な’, 文);
 ReplAll(‘思はねば’, ‘思◇ねば’, 文);
 ReplAll(‘思はれ’, ‘思◇れ’, 文);
 ReplAll(‘終はれ’, ‘終◇れ’, 文);
 ReplAll(‘終はる’, ‘終◇る’, 文);
 ReplAll(‘笑はれ’, ‘笑◇れ’, 文);
 ReplAll(‘住まはれ’, ‘住ま◇れ’, 文);
 ReplAll(‘給は’, ‘給◇’, 文);
 ReplAll(‘身がはり’, ‘身◆◇り’, 文);
 (非助詞「が」のマークもつけた)

2. 1. 2. が判定部

ここでの判定によるマーク付けは、後の採取部で取り上げられない「が」を意味する。すなわち、求める助詞「が」ではないというマーク付けである。助詞「が」を自動判定する一般的な規則は、菅沼明⁽²³⁾を参考にし、それを規則が(1)～規則が(4)にまとめた。保田序説からは別途規則が(5)を設けた。規則の適用は規則番号の順におこなった。

助詞「が」の区分については、益岡隆志⁽¹⁰⁾は{格助詞, 接続助詞}というように分類している。ここでは、益岡に従う。特に接続助詞については、将来文章論⁽¹³⁾⁽¹⁴⁾における文接続の関係に注目する意味から、特別に■マークを付与し、そのまま結果の表示に用いた。

規則が(1)

文中の「をが」における「が」は、助詞「が」ではないと考え、文中の総ての「をが」にマーク付けをする。ただし、「かをが」があれば、採用する。

ReplAll(‘をが’, ‘を◆’, 文); ReplAll(‘かを◆’, ‘かをが’, 文);

規則が(2)

文の先頭が「が」なら、助詞「が」ではないと考え、先頭の「が」だけをマーク付けする。

IF StrPos(‘が’,文)=先頭 THEN Repl(‘が’,‘◆’,文);

規則が (3)

文中の「が」の直後が「ん」撥音、ないし「つ」促音ならば、助詞「が」ではないと考え、文中の総ての「がん」「がつ」にマーク付けをする。ただし、保田序説においては、歴史的仮名遣いとしての促音は、「っ」ではなくて大文字使用である。ここでは、まぎらわしいものを別途に採用した。

ReplAll(‘がん’,‘◆ん’,文); ReplAll(‘がつ’,‘◆つ’,文);

ReplAll(‘◆つひに’,‘がつひに’,文); (採用)

ReplAll(‘◆つけら’,‘がつけら’,文); (採用)

規則が (4)

ここでは、接続助詞の判定をおこなった。接続助詞の可能性のあるものには、■マークを付け結果の表示に用いた。一般的には、格助詞「が」は名詞・連体形・助詞「の」等に接続し、接続助詞「が」は用言の終止形に接続するといわれている。なお「ところが」の場合には、次のような文例があり得るので、字面だけでの判定は困難である。これについては、今回は判定をしなかった。

金をもらったところが失った。

: 接続詞「ところが」

金をもらったところが宇治でした。

: 名詞「所」+格助詞「が」

ReplAll(‘うが’,‘う■’,文);

ReplAll(‘るが’,‘る■’,文);

ReplAll(‘くが’,‘く■’,文);

ReplAll(‘ぐが’,‘ぐ■’,文);

ReplAll(‘すが’,‘す■’,文);

ReplAll(‘ぶが’,‘ぶ■’,文);

ReplAll(‘す■た’,‘す◆た’,文);

ReplAll(‘いが’,‘い■’,文);

(「姿」を採用した)

ReplAll(‘だが’,‘だ■’,文);

ReplAll(‘つが’,‘つ■’,文);

ReplAll(‘たが’,‘た■’,文);

ReplAll(‘ひとつ■’,‘ひとつが’,文);

ReplAll(‘んが’,‘ん■’,文);

(「一つが」を採用した)

ReplAll(‘さん■’,‘さんが’,文);

ReplAll(‘ぬが’,‘ぬ■’,文);

(「～さんが」(敬称)を採用した)

ReplAll(‘むが’,‘む■’,文);

規則が (5)

保田序説で採取した助詞「が」ではない「が」。これに該当する文中の総ての「が」に、マーク付けをする。以下の用語には、歴史的仮名遣いと、保田自

身の文体の特徴が表れている。ひらかな遣いの多用は規則は（5）の解説と同じである。27例を規則とした。

ReplAll('あこがれ', 'あこ◆れ', 文);	ReplAll('やがて', 'や◆て', 文);
ReplAll('あがり', 'あ◆り', 文);	ReplAll('ゆがめ', 'ゆ◆め', 文);
ReplAll('あながち', 'あな◆ち', 文);	ReplAll('わけがな', 'わけ◆な', 文);
ReplAll('ありがた', 'あり◆た', 文);	(「訳がない」の慣用扱い)
ReplAll('かがやく', 'か◆やく', 文);	ReplAll('事がら', '事◆ら', 文);
ReplAll('かがやい', 'か◆やい', 文);	ReplAll('思ひがけ', '思ひ◆け', 文);
ReplAll('すてがた', 'すて◆た', 文);	ReplAll('おのが', 'おの◆', 文);
ReplAll('ちがひ', 'ち◆ひ', 文);	ReplAll('わが', 'わ◆', 文);
ReplAll('ちがふ', 'ち◆ふ', 文);	ReplAll('我が', '我◆', 文);
ReplAll('つがれ', 'つ◆れ', 文);	(以上3行は大野参照)
ReplAll('つなが', 'つな◆', 文);	ReplAll('見当がつ', '見当◆つ', 文);
ReplAll('ながら', 'な◆ら', 文);	ReplAll('甲斐が', '甲斐◆', 文);
ReplAll('ひがごと', 'ひ◆ごと', 文);	ReplAll('事が事', '事◆事', 文);
ReplAll('ひが言', 'ひ◆言', 文);	ReplAll('類がな', '類◆な', 文);
ReplAll('まがふ', 'ま◆ふ', 文);	(以上4行は、各々慣用扱い)

(漢字事例：「紛う」)

2. 1. 3. 採 取 部

ここでは、それぞれは判定部、が判定部を經由した1文に対して、最後の判定をする。すなわち、何らのマークも付けられていない「は」は、少なくとも「提題助詞は」あるいは「取り立て助詞は」であるものとして、「は」に替えて☆マークをつけ、結果の表示をする。おなじく、何らのマークも付けられていない「が」は、「格助詞が」であるものとして、「が」に替えて★マークをつけ、結果の表示をする。また、■マークはそのまま、接続助詞とみなして表示する。

2. 1. 4. 棄 却 部

ここでは、助詞「は」、助詞「が」を一つも含まない文だけを、そのまま表示する。

2. 1. 5. プログラム「はが文構造表示」の表示の説明

プログラム「はが文構造表示」による結果の一部を、図2にあげた。

```

.....
..... 満目冬景色の中に、
..... 敷松葉のくれなゐを見るやうな色★
..... あつたのだらう。
「 1: 8: 8」
かういふ色★
..... なければ、
..... 無といふものもないと思つてゐる。
「 1: 9: 9」
このうへもないもの、
このうへもない世界の中で、
おのれといふもの☆
..... 忘れられてしまふのである。
-----
^ 1: 10: 10 全く無なのである。
「 2: 1: 11」
そのころに☆
..... 早い年なら、
..... 愛宕のいただきに雪★
..... 見られ、
..... その雪★
.....
..... 真紅にか◆やく時★
..... ある。
「 2: 2: 12」
富士山のすばらしさ☆
.....
..... その山容の色彩★
..... 季節と時間によつて千変万化するときいてゐる。
「 2: 3: 13」
その片鱗☆
..... 見たこと★
..... ある■、
..... 私今生の願ひを一つ申せと言◇れるなら、
..... 富士山の眺められる富士川の畔に家居して、
..... この霊峰のいのちの粧ひを拝する日々を、
..... わ◆世の終瀬としたいと思つてゐたこと★
..... あつた。

```

図2 「はが文構造表示」による表示

「 1: 1: 1」
日本の文学史
序説
一

私★
……日本の美術史を書き了へたの☆
……、
……昭和四十二年の秋だった。

^ 1: 2: 2<その明け方の晩秋のけしきをよく覚えてゐる。

「 1: 3: 3」
二年に亙る著作をまづまづを◇つたといふ安堵感もあつたからだろうか、
その朝のわ◆山の家の風物☆
……、
……これ★
……今生かと思ふばかりのものだった。
「 1: 4: 4」
秋の黄葉☆
……、
……春の花より化粧賑かである。
「 1: 5: 5」
近づく冬の支度に、
いくらかつみ重なつた敷松葉☆
……、
……朝の露にぬれ、
……あけぼのの光りにあつて、
……くれないゑに匂つてゐる。
「 1: 6: 6」
冬の庭の敷松葉に、
朝日夕陽のさす時☆
……、
……真紅に近い色彩となる。
「 1: 7: 7」
芭蕉翁★
……自然の日本の風物の中で考へたわびさびの美観に☆

「 3: 1: 14」
旧著のことから書きだしたの☆
……、
私★
……日本の文学史を書かうと思つた因縁として、
……申さねばならぬ思ひ★
……あるからだ。
「 3: 2: 15」
その朝のわ◆家のかみの池や田邑（タムラ）のみささきの秋色の花やかさ、
この自然のおごりの、
いささかの滅びの予感もふくまぬ世界☆
……造形美術を以てして☆
……うつつこと★
……出来ない。

^ 3: 3: 16<仮りの相として、約束ごとのやうに抽象してみても、ま

とに心細い。
「 3: 4: 17」
俗世のしくみにまぎれこんでゐたら、
それでも、
絵だ、
美だ、
真だといつてをられよう■、
私★
……今立つてゐる世界で☆
……、
……もうそんな軽薄さ☆
……、
……まづおの◆心★
……第一番にうけ入れてくれぬ。

ここでは、その一例を図3として抽出し、説明を加える。

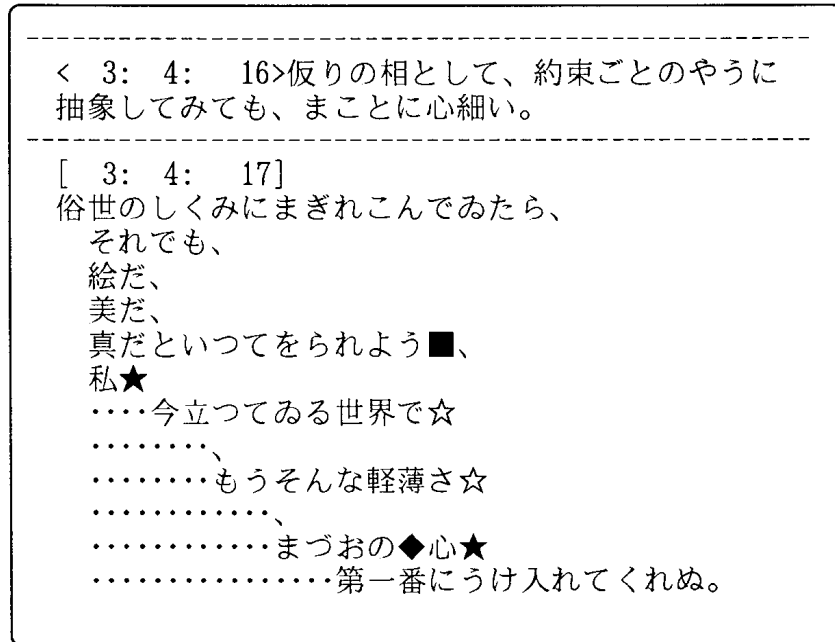


図3 表示の説明（☆助詞「は」、★格助詞「が」、
■接続助詞「が」、◆非取扱「が」）

図3で、最初の-----で囲まれた文（仮りの相として、…）は、助詞「は」と「が」を全く含まない文、すなわち棄却部で対象となった文を表す。またここで、〈 〉で括られた数字〈3:4:16〉は、全体で段落文番号を表している。最初の数字は文章先頭からの段落番号。次は段落内での文の順序番号、最後が文章先頭からの累積文番号を意味する。よって、これは〈第3段落、段落内第4文、通算第16文〉ということになる。

次に、[]以下の文は、助詞「は」ないし「が」を含んだ文章の、構造的な表示を意味する。[]で囲まれた数字の意味は、〈 〉と同様である。まず、冒頭の句（俗世のしくみにまぎれこんでゐたら、）は、それ以下の句よりも1文字分ずらせて強調してある。これは、時枝による文章論において、文章ないし文の冒頭に意味付けをおいた説に従った。この場合「句」とは、「,」までの部分で区切っている。

冒頭を除いた以下の「。」までに現れる句は、次のような規則によって表示

される。文中に「,」だけがあればそこで改行し、次の先頭位置は同じにする。文中に「☆ (助詞は)」ないし「★ (格助詞が)」があればそこで改行し、次の先頭位置は「…」を累加することにより、行頭を下げる。ただし、「☆,」と「★,」の場合には、そこでの後続句に対する論理的ないし感覚的な休止時間を表現する意味で、「----,」だけの文字列を次行に加えている。この図3の深部に含まれている文の構造を、擬似的に書き直したものが、次の図4である。

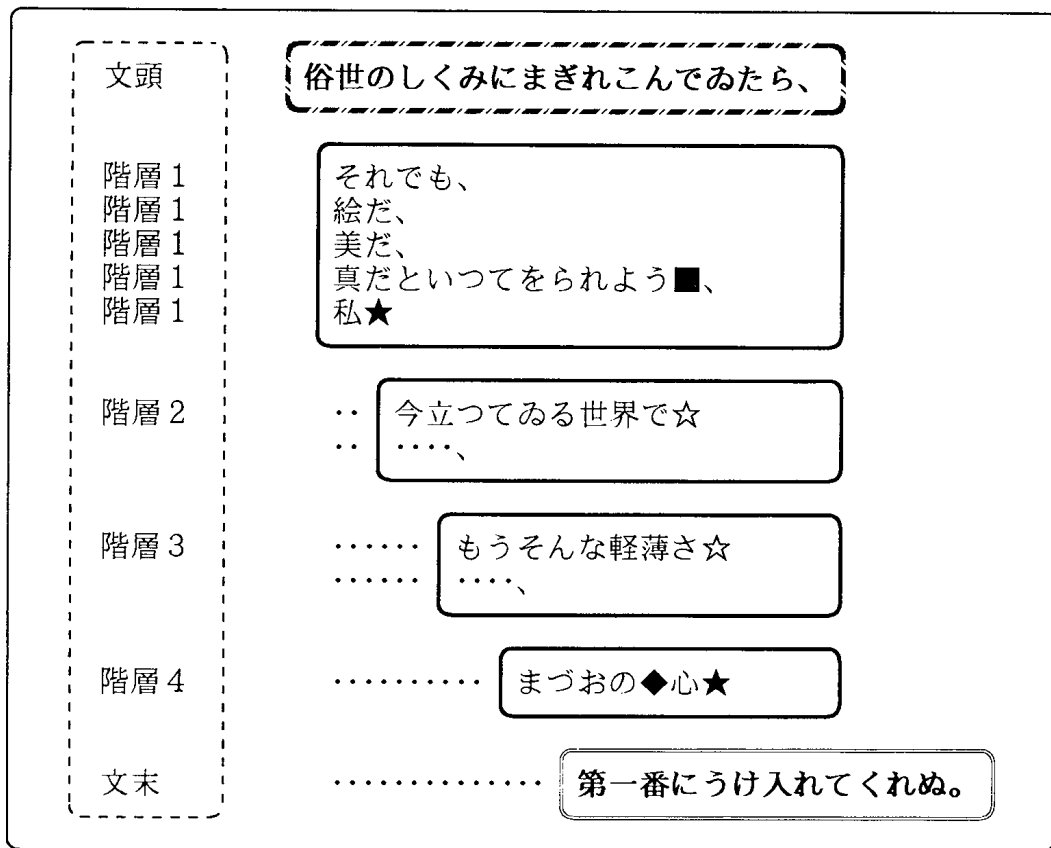


図4 「は」「が」「…、」による文の構造表現

この図4では階層1最後の「私★」に着目してみる。この「が」は直後の階層2「今立つてゐる」にだけ係っているという一般的な「が」の近距離係りが⁽²⁴⁾うなずけると同時に、「私」という言葉が強い力を持っていて、まるで遠距離係りの「は」のように、文末「第一番にうけ入れてくれぬ。」まで、影響しているような事実に気づく。⁽¹⁵⁾⁽²⁴⁾たしかにその係りを仮に1文にした「私が、第一番にうけ入れてくれぬ。」というのは助詞の誤用等による非文の典型ともいえる

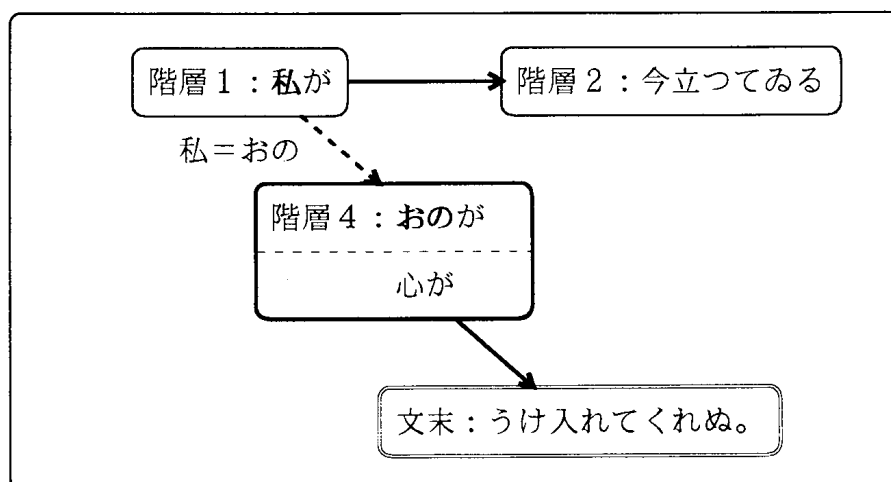


図5 「が」の深層係り構造

表4 「は」と「が」との統計

	助詞「は」	格助詞「が」	接続助詞「が」
出現数	236	103	25
段落平均	10.7	4.7	1.1
文平均	1.1	0.5	0.1
	非取扱「は」	非取扱「が」	
出現数	27	33	
段落平均	1.2	1.5	
文平均	0.1	0.2	

表3 段落と文の統計

段落	文数	段落平均文数
22	213	9.7文

が、文全体としてながめれば保田の文体がよく出ていて理解がしやすい。おそらくそれは、階層4の「おのが心が」が階層1の「私が」と同値になっていて、文の理解に錯綜を与えるぎりぎりのところで、文を形作っているのであろう。それは、図5のような深層の構造を持っていると解釈できる。プログラム「はが文構造表示」はこのような文ないし文章の解釈に視覚的な理解の補助を与えているといえよう。

2. 1. 6. プログラム「はが文構造表示」による表示の結果

プログラム「はが文構造表示」によって保田序説を分析した結果、この文章における文の統計的な数値は表3～4のようになった。結果は、悉皆調査とした。

第3章 「名詞+ {は | が}」の調査

ここでは保田序説から、「名詞+ {は | が}」を含む文をひろい、知識のネットワーク、すなわち知的ハイパーテキスト構築の可能性を探った。ここでの知識とは、1. 4. で定義したような、1文で明瞭な意味の完結したものをさす。

3. 1. 「は」「が」を伴った名詞の抽出

まず、「はが文構造表示」リストから、☆や★の直前の名詞が具体的な物や人、概念をさす場合、これを手技によって抽出した。ただし、人に関しては「大伴氏」「大伴の人々」「親房卿」「芭蕉翁」「平林英子さん」に見られるような、保田の文体に固有な言い回しは、これをそのまま採用した。別の[5:6:43]「大伴氏にうちかたうとした淡海公の家族ら★」は、ここでの手技による規則構成を越えているので、藤原不比等(淡海公)を抽出する事を保留した。このような保田與重郎にみられる韜晦は、不比等を淡海公と呼ぶようなところに多く見られる。たとえ古典の中では淡海公と呼ぶことが教養常識であったとしても、現代の読者には無理が生じる。これを通常の記事解析だけで解決することは困難であろう。ここに、専用の辞書の必要性が明白となるといえよう。

具体的には、先の抽出名詞の性格に加えて、以下のような簡単な規則も追加した。参考にしたのは、益岡⁽¹⁰⁾による名詞の解説部分である。

- (1) 「指示詞+ {☆ | ★}」は捨てる。「それは」「あれが」等。
- (2) 「助詞+ {☆ | ★}」は捨てる。「家には」「～とは」「～のが」等。
- (3) 「形式名詞+ {☆ | ★}」は捨てる。「～ことが」「～ものは」等。
- (4) 「数量名詞+ {☆ | ★}」は捨てる。「～ときは」等。
- (5) 「疑問を表す名詞+ {☆ | ★}」は捨てる。「誰が」等。

以上のような規則によって抽出した提題ないし主格部分をもつ句の一部を図

～	
[5: 4: 41]	当時の 大伴氏☆
[5: 5: 42]	大陸との外交通商の実権を掌握してみた 大伴の人々☆
[5: 9: 46]	……当時の世界中の珍貴な物産や 美術工芸品☆
[5:11: 48]	この 博物館☆
[5:17: 54]	日本人☆
[5:18: 55]	つまり 「源氏物語」☆
～	
[5: 2: 39]	唐土の 詩人★
[6: 3: 58]	……万葉集と 王朝文学★
[6: 4: 59]	……本居宣長翁の 学問★
[6: 7: 62]	……近世国学の文明論や 文芸論★
～	

図6 名詞+ {は | が} を伴った句の抽出 (一部)

表5 抽出した「名詞+ {は | が}」(67カ所 (☆41, ★26) : 異なり60種)

「源氏物語」☆	後水尾院★	大伴の人々☆	美術工芸品☆
「日本外史」☆	行事★	大伴氏☆	美術史☆
まつり☆	国力☆	朝廷★	美術史★
みち☆	山陽★	勅使★	風雅☆
英子さん★	思想★	鉄斎先生★	文学☆
王朝文学★	指導者☆	天皇☆	文芸論★
画壇☆	詩人★	天皇☆	文士☆
雅澄翁☆	字☆	天皇☆	文字☆
解釈☆	上人★	天皇陛下★	文章☆
外国人文士★	神々★	天子☆	文人儒者☆
革命☆	神道☆	道★	平林英子さん★
革命家☆	親房卿☆	日本人☆	理念☆
学芸☆	親房卿☆	日本人★	龍穴☆
学問★	親房卿★	日本文学☆	歴史家☆
教典★	政治☆	日本文学★	歴史書☆
契冲☆	太宰府★	芭蕉翁★	老荘☆
契冲阿闍梨★	体制☆	博物館☆	

6に表示した。☆をともなう部分は41例, ★を伴う部分は26例であった。今後このような「具体的な物や人, 概念」をさす名詞は, 保田の文章から手技で抽出することのある程度の範囲で継続し, それがまとまったときに見直し, 核となる「保田與重郎の辞書」とする方法が必要と思われる。なお図6は島探索プログラム⁽⁴⁾によって出力した結果を利用した。図中の「||」は, 抽出した名詞のための先頭区切り記号である。図6の名詞だけを整理したものが表5である。

3. 2. 言葉同士の結びつき（リンク）をモデル化する

これらの名詞をどのように結びつけ、知識ネットワークに構築していくかのモデルを次に説明する。例として、同じ言葉にそれぞれ「は」と「が」が付いた北畠親房を取り上げてみる。出現箇所は以下の3カ所である。

[16： 1：154] 北畠の || 親房卿★

[16： 4：157] || 親房卿☆

[17： 3：162] || 親房卿☆

なお、以下では事例中総ての「私」は保田與重郎に置き換えて説明する。このことによって、評論家としての保田與重郎の対象化を図った。また例図には、段落を越えない範囲で前後文を添えた。

図7からは、「北畠の親房卿★「神皇正統記」を著されたの☆」という知識が直接得られる。これは★であるから、一般的には直後の述部だけにかかるという考えをそのまま適用できる例である。この部分をそのまま蓄積し、表示するだけで知識の伝達は可能である。また、他の言葉と結びつける手法としては、目的語となる具体性を持った名詞**神皇正統記**を親房卿に直接の関連をもった重要語として抽出し、後述の図9、図11に結合することができる。保田序説では、**親房卿★「神皇正統記」**はあっても、**親房卿☆「神皇正統記」**はないから、北畠親房とその著作を結びつける{は|が}による表現はここだけである。しかし、その1カ所でもあれば、文中の多くの間連語に結合させることが可能となる。後述の図11では、「日本外史」との関連が生じている。

同図7の「**親房卿☆**一部の史料ももたないで歴史の書を著述された。」は明瞭な知識を文として表現しているとはいいがたい例となっている。この部分だけからなにかを得る読み手は高度な知識をすでに持っているといえる。実際には他の書で、親房は『皇代記』を手元に置いていたとの言及がある。⁽²⁵⁾「は」は、後続文に対して遠距離の係りを持つ場合もあるが、文末までに限定した本稿での規則の場合、後文はあくまで参照部分に過ぎない。よって、この例は単純な1文によって明瞭な知識事例を提供するという本稿の主旨では、範囲を外れている。

【親房卿★】 【親房卿☆】 段落内前後行
[16: 1: 154] 北畠の親房卿★ ……「神皇正統記」を著されたの☆ ……、 ………関東僻地の孤城で、 ………四面みな敵の重囲下にあつた。
[16: 2: 155] ただ絶体絶命の敗滅★ ……その運勢のやうに思ひれた。
[16: 3: 156] 出撃☆ ……狂者のなすところ、 ……逃走☆ ………愚者の計だつた。
[16: 4: 157] さういふ死を待つやうな絶対の場に当つて、 親房卿☆ ……一部の史料ももたないで歴史の書を著述された。
[16: 5: 158] これ★ ……まことの歴史の書として、 ……日本人★ ………かつて著した無数の書籍の中でも、 ………かけかへのない貴い文学であつた■、 ………その志を思ふ時、 ………同じ民族の一人として、 ………保田與重郎☆ ………耐へ難い感動にうたれる。

図7 【親房卿★】 【親房卿☆】 段落内前後行

図8での「親房卿☆実践して教へられた。」も先述の「は」と同じである。この文だけでは明瞭な知識を提供しているとはいいがたい。ただし文内を遡及する論理が組み込めれば、「教へられた」の内容から保田與重郎が北畠親房をいわゆる「道義の人」としてとらえているという知識は伝達されるであろう。本稿の考え方では、このような場合着目文の前後文を提示することで解決を図っている。

次の図9からは、☆ないし★をともなった具体的な名詞と、それに直接的に関連した具体的な名詞が、文章の他の部分でどのように扱われているかを見るための、事例参照の提示とする。

【親房卿☆】段落内前後行
[17: 2: 161] 敗亡なすすべ★ ……ない。
[17: 3: 162] さういふ時に当つて、 道義と人倫を未来に恢弘するためになしうる唯一の方法を、 親房卿☆ ……実践して教へられた。
[17: 4: 163] 保田與重郎☆ ……このことを終戦の直後にくりかへし思ひかへし、 ……獄中で自国の歴史を著述したアジア近來の 革命名士のことをあ◇せて思ひ起した。

図8 【親房卿☆】段落内前後行

【親房卿】【神皇正統記】段落内前後行
[14: 3: 130] それらの文章☆ ……何編か出来て、 ……今も 保田與重郎☆ ……それに満足してゐるのである。
[14: 4: 131] 以前 保田與重郎☆ …… 親房卿 の 神皇正統記 の思想を批評したこと★ ……あつた。
----- < 14: 5: 132 >又泪をこらへてその一節を幾度も 誦したこともあつた。 -----

図9 【親房卿】【神皇正統記】段落内前後行

同図9では、「**保田與重郎☆親房卿**の**神皇正統記**の思想を批評したこと★あつた。」という、非常に明瞭な知識が得られる。後続の第132文「又」以下によって、読者はそれが親房批判であったことが推量できる。しかし、「親房卿」で結びつく図10の145文、「神皇正統記」で結びつく図11の178文によって、今は保田が批判的ではないことも参照できる。

図10からは、保田與重郎の初期「北畠親房批判」への緩解が見られる。

【親房卿】段落内前後行
[15: 1: 145] 保田與重郎☆ ……親房卿に玉子をぶちつけて、 ……そして泣いてみたのである。

< 15: 2: 146 >ずみ分遠い年少客気の日の思ひ出である。

図10 【親房卿】段落内前後行

【神皇正統記】段落内前後行
[19: 5: 177] 山陽★ ……外史をかいて、 ……多数の参考書目をかかげた■、 ……本文に☆ ……一行の銜学もない。
[19: 6: 178] しかし「日本外史」☆ ……「神皇正統記」にくらべると、 ……我◆両手でさげられる程の石のやうである。

図11 【神皇正統記】段落内前後行

図11を参照して得られる知識は、「日本外史」に比べて「神皇正統記」の、保田自身における重さを推量することができる。すなわち『我◆両手でさげられる程の石のやうである。』という表現の強さに保田の考えが込められている。

最後の図12は図7～図11にみられる知識のネットワークをまとめたものである。以上のようなモデルを通して、初歩的な「自動的ハイパーテキスト化」ならびに「知的ハイパーテキスト化」の可能性が生まれる。

3. 3. 「保田與重郎☆」と「保田與重郎★」

文中での「私」だけに着目し、「は」「が」の出現を調査した。表示は3. 2. と同じく「私→保田與重郎」とした。結果は、「保田與重郎☆」は52例/236用例中、「保田與重郎★」は5例/103用例中であった。以下では、抽出された文に対しての考察を行う。

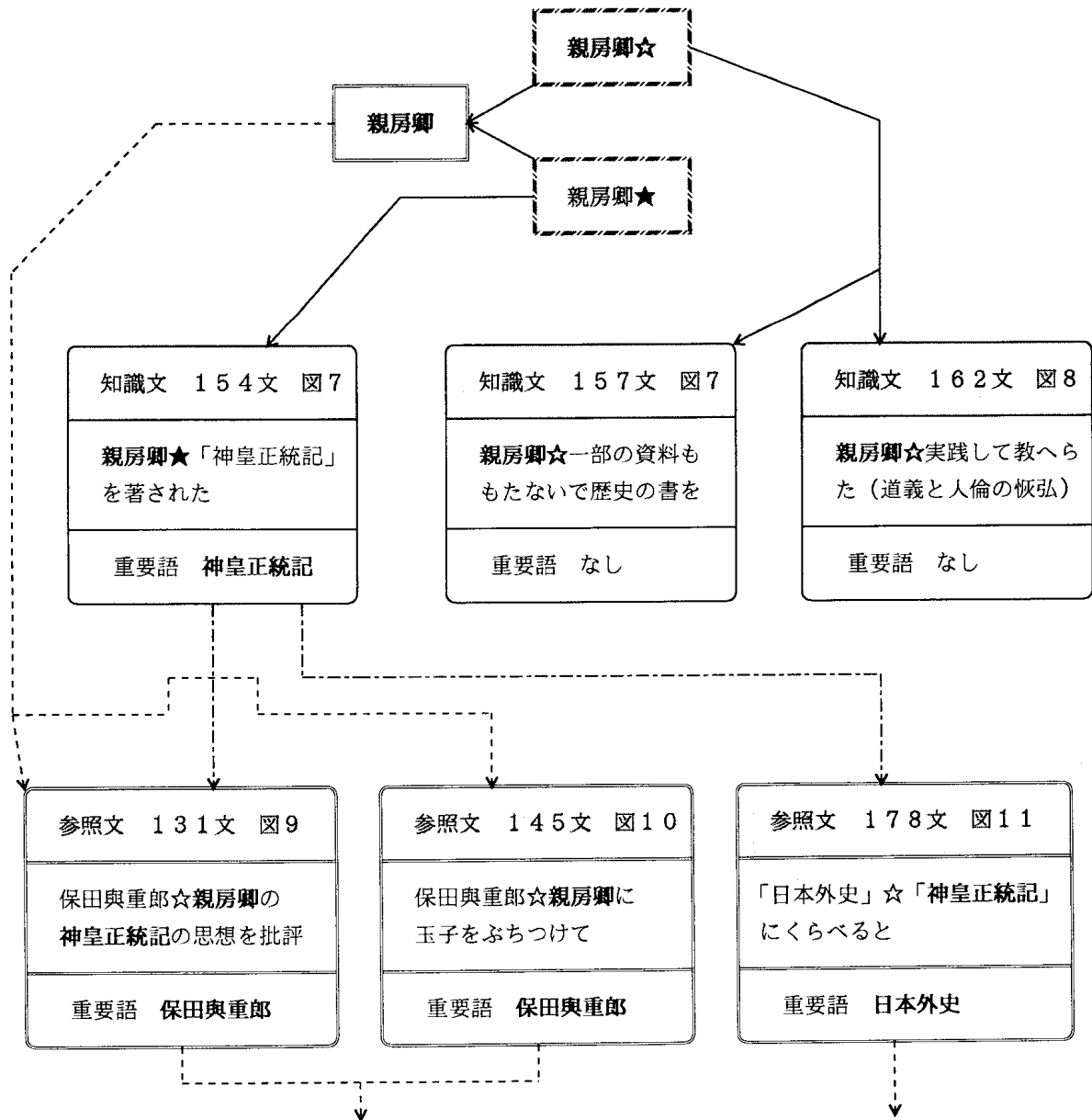


図12 「親房卿+ {は | が}」に着目した知識のネットワーク

3. 3. 1. 保田與重郎☆ (「私は」という表記)

文中「私は」という文は52例あった。次にその中からわかりやすい知識を提示できるものを選んだ。結果は15例あった。これは15/52/236という比率になり、全用例中の6%強である。全213文数中では、文の7%に比較的に意味を抽出しやすい「私は」があるとなる。知識抽出の規則は、述部に重要な名詞があり、その文だけで知識が得られるものとした。また各事例下部に短縮した

文を添えた。それに該当しない例は後で解説した。

- [6 : 3 : 58] 保田與重郎☆中学生の時代に万葉集をよむ時，土佐の鹿持雅澄翁の「万葉集古義」に学んだので，万葉集と王朝文学が一つだつたところをまづ教はつたのは全く幸ひだつた。

保田與重郎☆ 鹿持雅澄翁の「万葉集古義」に学んだ

- [9 : 6 : 92] 保田與重郎☆伝教大師を文学史で多少知つてゐたから，一乗寺の御影にあこがれることなみなみでなかつたのである。

保田與重郎☆ 伝教大師を知つてゐた

- [10 : 8 : 101] 保田與重郎☆年少の日から，自分の文学の立場を，後鳥羽院以後の隠遁詩人の志といふものにおいてきた。

保田與重郎☆ (文学の立場を,) 後鳥羽院以後の隠遁詩人の志においてきた

- [11 : 1 : 109] 保田與重郎☆日本の美術史をかきすすんで，近古近世に及んだ時，わが執心の本意をいふのに，文学史でなければならぬと感じたのである。

保田與重郎☆ 文学史でなければならぬと感じた

- [12 : 1 : 110] 終戦の直後の保田與重郎☆，なすことも少いままに，神々が地上を遊幸されてゐたころの歴史を考へてゐた。

保田與重郎☆ 歴史を考へてゐた

- [12 : 3 : 112] 保田與重郎☆人間の世界の歴史のまへの，神々の国の地理や風景を描き出し，さういふものの二三は文章にし，それは極く僅の人のところをひいたことを知らされた。

保田與重郎☆ (人間の世界の歴史のまへの,) 神々の国の地理や風景を描き

- [13 : 1 : 113] 何をするといふこともない日々に，保田與重郎☆戦争中の徴用によつて荒廃させられた水田を復元して，そこへ米を植ゑた。

保田與重郎☆ (水田を復元して,) そこへ米を植ゑた

- [14 : 4 : 131] 以前保田與重郎☆親房卿の神皇正統記の思想を批評したことがあつた。

保田與重郎☆ 親房卿の神皇正統記の思想を批評した

- [14 : 17 : 144] これを逃避といふなら，保田與重郎☆「コギト」や「日本浪漫派」のむかしから，逃避の道一筋を危く渡つてきたやうなものだ。

保田與重郎☆ 逃避の道一筋を危く渡つてきた

- [16 : 5 : 158] これがまことの歴史の書として、日本人がかつて著した無数の書籍の中でも、かけかへのない貴い文学であつたが、その志を思ふ時、同じ民族の一人として、保田與重郎☆耐へ難い感動にうたれる。
保田與重郎☆ 耐へ難い感動にうたれる
- [21 : 1 : 185] 保田與重郎☆ 現今の歴史研究の方法や、歴史論文の方法をすべて無意味とはいはないが、大体無関心と断言する。
保田與重郎☆ (歴史論文の方法を) 大体無関心と断言する
- [22 : 1 : 197] 保田與重郎☆終戦後は、さういふ歴史書の構成を考へ、時々昂奮することもあつたが、それも概して懶惰を日常としてゐた。
保田與重郎☆ 懶惰を日常としてゐた
- [22 : 8 : 204] 日本の美術史が近世にすすんで、私のところはゆきなづみ、保田與重郎☆齊藤氏にその嘆息をもらした。
保田與重郎☆ 齊藤氏にその嘆息をもらした。
- [22 : 11 : 207] 保田與重郎☆日本の文学史では、近世を旨として、文人の志の系譜をたどりたい。
保田與重郎☆ 文人の志の系譜をたどりたい。
- [22 : 12 : 208] 「日本文学」といふべきものは、後鳥羽院まで、さういふ日本文学が、どういふ形で国の人々の心に生き、国の人々のところを生かしたかといふ、二つの事理のやうな、しかも大本は一つのやうな、そんな文学史を保田與重郎☆かいてみたいと思つた。
保田與重郎☆ (そんな文学史を) かいてみたいと思つた。

比較的に意味がわかりやすい文の抽出とはいっても、次のような例は文末における文の様態を詳しく分析しないと、誤った情報を知識として与える。

- [14 : 17 : 144] これを逃避といふなら、保田與重郎☆「コギト」や「日本浪漫派」のむかしから、逃避の道一筋を危く渡つてきたやうなものだ。

この「～やうなものだ。」の部分解析する文や文章の高度な自動化は未だ考慮をしてはいない。これはモダリティに関わることである。ここでは、「私は」が、保田與重郎の関わった文学運動である「コギト」「日本浪漫派」と同一文中に出現しているという事実止めておく。しかしそれだけでも明瞭な知識リンクの可能性を持っている。

この「保田與重郎☆」の部分から、文だけでは明瞭な知識を受け取れない部

分の例を次に説明する。それは対象となった全22段中の中ごろに位置する第13段の文であり、知識としては取り上げなかった。

第13段／全22段

[13： 3：115] 以前の保田與重郎☆文章や文字は、空にかけるものであることを知らなかつた。

[13： 4：116] 保田與重郎☆そのころさういふことを発明したのである。

[13： 8：120] 保田與重郎☆非常に閑があつたので、空に字をかいて手習ひをしてゐた。

[13：12：124] その時目をつむつてかくといふことに保田與重郎☆気づかなかつた。

[13：14：126] 保田與重郎☆青い空に小さい字をかいてゐた。

この部分は、日本語文としては成り立っていても、「[13:14:126] 私は青い空に小さい字をかいてゐた。」という要約から、本稿で限定的に用いている知識を導き出すのは容易ではない。これはむしろ詩の文章といえよう。この5文、あるいは第13段全部を読んだだけで、知識を得ることができたと思う人はまずいないであろう。しかし、この文章は、戦後まもなく出た『日本に祈る』⁽²⁶⁾を読んだ後ならば、保田與重郎にとって非常に重大な精神状態の過程を述べた部分として、その位置づけの深さを知ることができる。要するに、終戦後自ら筆を断ち、また公職追放などにより筆を断たれた保田が、それでもなお、後鳥羽院以後隠遁詩人の系譜を信じ、紙ではなく空に文字を書いたという事実が背景にある。その書の自序部分は、4頁にわたり全文カタカナで書かれ、印象としては神語に近いものである。

保田與重郎の全ての文章がコンピュータ上にあつたとして、この5文から、あるいは、第13段全体から『日本に祈る』に結合の糸（リンク）が張られていても、現在の本稿での規則や技術ではこれを明瞭な知識として抽出することはおぼつかない。それには文章の感性の部分をうまく扱う考え方をとりいれなければならないのであろう。これは単純な知識命題抽出の扱いを越えている。よって、「保田與重郎☆」からは外した。

3. 3. 2. 保田與重郎★（「私が」という表記）

文中の「私が」をひろいだしてみると次のように5例だけであり、103用例

中 5 %弱。全文数213文からみれば 2 %強の割合である。

- [1 : 1 : 1] 保田與重郎★日本の美術史を書き了へたのは、昭和四十二年の秋だった。
- [3 : 1 : 14] 旧著のことから書きだしたのは、保田與重郎★日本の文学史を書かうと思つた因縁として、申さねばならぬ思ひがあるからだ。
- [3 : 4 : 17] 俗世のしくみにまぎれこんでゐたら、それでも、絵だ、美だ、真だといつてをられようが、保田與重郎★今立つてゐる世界では、もうそんな軽薄さは、まづおのが心が第一番にうけ入れてくれぬ。
- [9 : 5 : 91] 保田與重郎★日本の美術史を著述しつつ、時々嘆息したのは、かういふ点だった。
- [20 : 2 : 180] 保田與重郎★文人だからかく申すのである。

この5例の「保田與重郎★」をしてみると、いずれも「ほかならぬ私が」という意味で使われている。そのうち3例は「保田與重郎★日本の美術史を」「保田與重郎★日本の文学史を」「保田與重郎★日本の美術史を」というように、自身の著作を格助詞「を」で添えている。このことは、単に著作者であるという事実認定であるに過ぎないともいえる。しかし文芸評論家であることを越えて、日本の精神文明に対する思想家としての彼の位置づけを考えるなら、「私が」と語り出すときの保田與重郎の言葉は、揚言に近いといえよう。それは次のような強調を意味している。

(ほかならぬ) 私が (ほかならぬ) 日本の美術史や文学史を書いた。

また『日本の美術史』⁽²⁷⁾『日本の文學史』という二つの著作名は、文人のありようからなのか、「」なしで地の文に溶かせてはいるが、それは逆に普遍的な「日本の」という形容を際立たせてもいる。「日本の美術史」ならば単に書名に過ぎないが、「」を付けないと、日本の地文の一部であるような印象を与える。他の例では「神皇正統記」「源氏物語」「日本外史」ともに「」付きである。

残りの2例では、「保田與重郎★今立つてゐる世界」「保田與重郎★文人」では、「今立つてゐる世界」とはすなわち、彼の精神世界であり、「文人」というのは、後鳥羽院以後隱遁詩人の脈絡の中での、直言に近い意志表示である。こ

の2例は立場の宣言である。思想家にとっての著述が意志表示であるとするならば、前者の3例と合わせて、保田與重郎における「私が」というのは、非常に明瞭な「思想家ないし文人としての意志の現れ」であるといえよう。これを保田與重郎の語彙にあてはめると、「述志」ということになる。

わずか5例での推測は推測にすぎないが、ここでの結論は「私が」とあれば、それを含む文は、知識として、保田與重郎の「述志」部分である可能性があると、推量しておく。

第4章 結 論

はじめにと第1章では、言語工学と日本語の文法学の成果を図書館学に取り入れることの必要性と、本稿の背景を説明した。

第2章では、助詞「は」と助詞「が」を自動的に判別し、文の構造を明瞭に表示するための「はが文構造表示プログラム」を作成した。このための規則として「は」「が」それぞれに4つの一般的な規則を情報工学の成果から引用し、別に5つめの保田與重郎に固有な規則を設けた。結果として、「は」も「が」も、それぞれ第5の規則が必要になったのは、保田與重郎の場合、歴史的かな遣いの問題よりも、特殊な文体からくる用語の判別にむつきしさのあることが明確になった。このことから、保田與重郎の専用辞書の構築が必要となる。

第3章では、まず「は」と「が」を他の助詞等との結合を持たないもっとも単純な形でとりあげ、その直前に位置する比較的意味の明瞭な名詞の抽出を手技でおこなった。その結果「名詞+は」は41/236用例中で17%、「名詞+が」は26/103用例中で25%となった。これらの名詞を媒介にして、他の名詞との共起を調べることにより、知識の網を作り出していくモデルを、北畠親房を例にして提出した。これは、今後自動的にリンク網を構築するための準備であった。

次に、保田與重郎の場合、「私は」「私が」の「私」を保田與重郎に置き換えて、知識として抽出し提示すると有効であることがわかった。結果として「私は」の提題から得られる述部における明瞭な知識文は、15/236という比率に

なり、全用例中の6%強。全文数213文の7%であった。一方、「私が」については、103用例中5%弱。全文数213文の2%強であった。用例が5カ所であり、即断はできないが、その部分が結果として宣言になっていることがわかった。これは、述志という直言である。直言はその論理と感性の結果や経過を問わず、著述者の生の精神の表出であると推量できる。これは少なくとも保田與重郎に関しての知識として、もっとも重要な部分であるといえる。

それぞれから抽出できた割合は少量であるが、質の問題として考えたとき、あるいは大量のテキスト情報からなにを限定した知識として選定するのかを考えたとき、今後これらの手続きの自動化の意味は十分あると考えるものである。

以上の2章と3章とにおける助詞や名詞の判別には特別な形態素解析システムは使用していない。本稿での対象文の特殊性から考えて、あえて地の文に直接⁽²⁸⁾接触れることに重きをおいたからである。しかし、すでにJUMANなどの汎用性を持った有力な形態素解析システムがあり、利用者毎に文法の定義や単語間の接続関係の定義ができるようになってきた。今後、大規模な文章解析には利用してみたいと考えている。

おわりに

本稿で得られた結果をもとにして、今後の自動索引、自動抄録、自動的ハイパーテキスト化、自動分類等に結びつけていきたいと考える。⁽²⁹⁾⁽³⁰⁾難解といわれる保田與重郎の思想に対する基盤的な書誌を作成することがねらいである。その特殊性のなかで得られる、自然言語処理に必要とされる諸規則を、他の一般的な文章のそれと比較することにより、「日本語文章からの知識の抽出」をより普遍的な姿に近づけていきたい。

参考文献と注記

- 本稿で使用した開発環境は、NEC PC-9821ならびにFM-Towns上における、ボーランド社の一連のTurbo-PASCAL、Borland-Cである。また、保田與重郎の著作は総て旧版を使用し、本文の仮名遣いは残し、旧漢字だけを新漢字にあらためた。
- (1) 谷口敏夫「ハイパーライブラリー」『現代の図書館』31(2), 1993.

- p. 96-102。
- (2) 保田與重郎『日本の文學史』新潮社, 昭和47年. 序説は p. 7~18。
 - (3) 野田尚史『はとが』くろしお出版, 1985。
 - (4) 谷口敏夫「新しい全文検索アルゴリズム: 島探索(アイランドサーチ)方式」『情報の科学と技術』42(8), 1992. p. 727-731。
 - (5) 安本美典『文章心理学入門』誠信書房, 昭和40年。
 - (6) 長尾真『言語工学』昭晃堂, 昭和58年。
 - (7) 長尾真『画像と言語の認識工学』コロナ社, 1989。
 - (8) 黒橋禎夫, 長尾真他「専門用語辞典の自動的ハイパーテキスト化の方法」『人工知能学会誌』7(2), 1992. p. 336-345。
 - (9) 長尾真『人工知能と人間』岩波書店, 1992。
 - (10) 益岡隆志, 田窪行則『基礎日本語文法 一改訂版一』くろしお出版, 1992。
 - (11) 益岡隆志『命題の文法』くろしお出版, 1987。
 - (12) 益岡隆志『モダリティの文法』くろしお出版, 1991。
 - (13) 時枝誠記『日本文法: 口語篇』岩波書店, 1950。
 - (14) 永野賢『文章論総説』朝倉書店, 1986。
 - (15) 大野晋『係り結びの研究』岩波書店, 1993。
 - (16) 大野晋『日本語の文法を考える』岩波書店, 1978。
 - (17) 足立巻一『やちまた上下』河出書房新社, 1990。
 - (18) 西田龍雄『西夏語研究を顧みて』京都大学文学部言語学教室, 平成4年。
 - (19) 谷口敏夫「2001年, フルテキストデータベースの未来」『情報の科学と技術』42(6), 1992. p. 498-506。
 - (20) 伊井春樹「『源氏物語』の読者たち」『新装版図説日本の古典7 源氏物語』集英社, 1988. p. 168-179。
 - (21) 菅沼明, 牛島和夫「日本語文章推敲支援ツール『推敲』におけるとりたて詞「は」の抽出法とその評価」『情報処理学会論文誌』32(11), 1991. p. 1392-1400。
 - (22) 保田與重郎『現代畸人傳』新潮社, 昭和39年。
 - (23) 菅沼明他「日本語文章推敲支援ツール『推敲』における字面解析手法とその評価」『自然言語処理(情報処理学会)』68-8, 1988。
 - (24) 野田尚史「複文における「は」と「が」の係り方」『日本語学』5, 1986. p. 31-43。
 - (25) 保田與重郎『日本語録』新潮社, 昭和17年。
 - (26) 保田與重郎『日本に祈る』祖国社, 昭和25年. 自序からの引用を一部記しておく。“来ルモノヲ引寄セテ眩ク声ハ, 余ニシカト聞エクルゾ。誰ニモワカラヌ言

葉デ、必ズワカル如クニ、汝ハ眩ク。紙無ケレバ、土ニ書カン。空ニモ書カン。”
この部分は、「磯田光一『比較転向論序説』勁草書房，昭和43年。」第4章「ナシ
ョナリズムの美学」にも部分的に引用され論評されている。

- (27) 保田與重郎『日本の美術史』新潮社，昭和43年。
- (28) 松本祐治，黒橋禎夫，宇津呂武仁，妙木裕，長尾真『日本語形態素解析システム JUMAN 使用説明書』京都大学工学部長尾研究室，1993。
- (29) 橋川文三『増補日本浪漫派批判序説』未来社，1965。この「イロニイと文体」では，保田の文体の特殊性（難解さ）が次のように要約されている。

“保田の文体を批判することは特別に重要である。それは一般に文体と思想が不離のものであるという以外に，保田がある時期に殆ど呪術的な魅力をもって一世を風靡したことには，その文体の持つ性格が与って大きかったからである。それはたしかに異様な文章であった。高見順が初期の保田の文書を見てめくらめくような印象を受け，自分たちの時代は去ったのかと嘆息したことを平野謙がつたえているが，それはまさに私たちが見たこともなく，これからみることもないような文章であった。”

- (30) 篠田一士『樹樹皆秋色』筑摩書房，1989。この「保田與重郎」では，保田與重郎の難解さが文体の変化をみることによって，ときほぐされている。