国立国語研究所学術情報リポジトリ

曖昧性解消過程解明のための多義語の分析

メタデータ	言語: Japanese 出版者: 公開日: 2021-06-11 キーワード (Ja): キーワード (En): 作成者: 柏野, 和佳子 メールアドレス: 所属:
URL	https://doi.org/10.15084/00003304

曖昧性解消過程解明のための多義語の分析

言語体系研究部　第二研究室　　柏野和佳子

(E-mail: waka@kokken.go.jp)

要旨：国語学の領域において，多義語の曖昧性解消過程は十分に解明されていない。他方，日本語処理においては，語の曖昧性を解消するための計算機用の辞書研究が進められている。曖昧性の解消に有効な辞書情報として，格パターン情報が知られていた。しかしながら，格パターン情報だけでは十分とは言えなかった。コロケーションの曖昧性の分析を通じ，曖昧性の解消には，名詞句，格助詞，述語という単純な格パターンだけではなく，連体修飾句，連用修飾句，格表示に関しての詳細な情報を盛り込んだ統語情報が必要であること，さらに，形態や文脈に関する情報も必要になることが分かった。今後，その必要な情報の内容を明らかにするために，多義動詞の曖昧性解消過程を解明する研究を行っていく。

キーワード：曖昧性の解消，多義語，コロケーション，辞書，日本語処理

1．はじめに

言葉の意味・用法についての研究は数多くあるが，複数の意味・用法をもつ，いわゆる多義語の曖昧性を解消する過程は，十分に明らかにされているとは言えないのが現状である。しかし近年，実際の文章や談話で使用されている語の意味が，何を手掛かりにしてどのように特定されるかを解明する重要性が認識されつつある。

本稿では，最初に，多義語の曖昧性解消過程の解明が求められる背景を説明する。続いて，これまで日本語処理研究において進められてきた，計算機による曖昧性解消のための辞書記述の概要を述べ，問題点を指摘する。最後に，現在取り組みはじめた，多義動詞の曖昧性解消過程解明のための研究概要について報告する。

2．背景

国語学の領域において，多義語の曖昧性解消過程は十分に解明されていない。その一方で，計算機による日本語処理研究においては，実用性の高い処理を行う上で多義語の曖昧性が大きな支障の一つとなっている。

例えば，計算機で日本語から英語に自動翻訳しようとする場合，日本語では同じ語であっても，別々の言葉に訳し分けなければならないということが多くある。このような時，日本語に現れる多義語の曖昧性を解消する必要が出てくる。人間は，すでに持っている知識を用いて推論することによって曖昧性を解消できるものであるが，計算機に同じように曖昧性を解消させるためには，言語に関するさまざまな情報を，あらかじめ辞書に記載し，持たせておかなくてはならない[柏野・本多98a，柏野・本多98b]。その情報を明らかにするためにも，国語学における多義語の曖昧性解消過程解明の研究が強く求められている。

3．計算機用辞書記述の実際と問題点

3.1　格パターン情報

計算機による日本語処理のための辞書としてこれまでに作成されたものには，ＥＤＲ辞書[ＥＤＲ93]，ＮＴＴ日本語語彙体系[ＮＴＴ97]，及び，筆者もその開発プロジェクトに参加していた，ＩＰＡＬ[ＩＰＡ97]がある。いずれの辞書にも語のもつ曖昧性を解消する情報として，格パターン，あるいは，格フレーム[Fillmore75]などと呼ばれるものが記載されている。これは，どのような意味素性をもつ名詞句が，どのような格要素を介して述語（動詞／形容詞／名詞＋ダ）と結びつくかを示すものである。

例えば，ＩＰＡＬ動詞辞書には「割れる」という動詞について，図１に示すように，２つの意味それぞれに格パターンと，慣用表現とが記載されている。ここで，図１の格パターン情報を利用して曖昧性を解消する例を一つ示そう。例えば，「鍋が割れる」という句が現れた場合，その「割れる」が，図１に示した，①，②の意味か，あるいは，慣用表現なのかを考えるとする。「鍋」という名詞は，①，②の格パターンに

割れる：
①一体であった物が粉々に分かれた状態になる。
【格パターン】
(N1[PHE/ACT]デ) N2[CON]ガ 割れる
N1 PHE:地震 / ACT:衝突
N2 CON:ガラス，茶碗，皿，氷，卵
【文例】地震で ガラスが 割れた。
【慣用】頭が～ように痛い。～ような拍手。
②一つのものがいくつかに分かれる。
【格パターン】
(N1[ACT/ORG]デ) N2[ORG/ABS]ガ
(N3[QUA]ニ) 割れる
N1 ORG:会議 / ABS:委員会
N2 ORG:党，組合 / ABS:意見
N3 QUA:二派，三つ
【文例】会議で 意見が 三つに 割れた。
【慣用】票が～。犯人・ホシが～。尻が～。
＊英文字３字は，意味素性を表す。意味素性とは，結びつく述語によって焦点が当てられる意味的な側面である。
PHE:自然現象，ACT:行為，CON:具対物，
ORG:組織，ABS:抽象物，QUA:数量

図１：ＩＰＡＬ「割れる」の記述

も，慣用表現にも記載されていない。慣用表現は名詞句が一致する場合に限られるので，この可能性が最初に排除される。①の意味か②の意味かを判別するためには，「鍋」の意味素性を判定する必要がある。実は，「鍋」も多義語であり，曖昧性がある。ＩＰＡＬ名詞辞書には，「鍋」について，図２に示すように，意味素性や結びつく述語例などが記載されている。これにより，「鍋」の①に具対物を表す意味素性[CON]があり，②に可食物を表す[EDI]のあることが分かる。さらに，意味素性[CON]で結びつく述語例には，「割れる」という語そのものが挙げられている。

鍋：
①食物を加熱調理するための底の深い器。
【意味素性】CON
【述語】割れる，熱くなる，へこむ，洗う
【文例】鍋に煮物が入っています。
②「なべ①」に様々な具を入れ，煮ながら食べる料理。
【意味素性】EDI
【述語】食べる，味わう，おいしい，うまい
【文例】みんなで鍋をつつく。
＊EDI:可食物

図２：ＩＰＡＬ「鍋」の記述

一方，図１の「割れる」の記述には，①に，[CON]という意味素性をもつ名詞句例との結びつきが示されているが，[EDI]についての記載はない。これらのことから，「鍋が割れる」という時は，物[CON]が壊れる意味の「割れる」①に，物[CON]を表す器の意味の「鍋」①が結びつく，ということが分かり，「割れる」と「鍋」両方の曖昧性が解消される。

3.2 コロケーションの曖昧性という問題

ところが，格パターンの記述を，名詞と述語，一語ずつの組み合わせだけに限ると，同一の名詞と述語の結びつきであるコロケーションの曖昧性が解消できないという問題がある。例えば，「骨が折れる」という時，「骨」という語に①人間の骨を指す意味と，②傘などの骨を指す意味とがあり，さらに句全体で慣用的に③「苦労する」という意味を指す場合もある。そのため，名詞と述語の結びつきだけでは，いずれの意味か特定することができない。

この「骨が折れる」のような場合は，各語にある曖昧性の解消に先立ち，コロケーション全体にある曖昧性を解消する方法を検討する必要がある。これまでの計算機用辞書では，このような，コロケーションのレベルで存在する曖昧性には十分に対処できていないという問題点があった。

3.3 コロケーションの曖昧性の分析

そこで筆者は，コロケーションの曖昧性の問題について考えるために，先に，曖昧性をもつコロケーションの用例（延べ10,415例，異なり4,578例）を定量的に分析した（詳細は[桑畑他98]に報告している）。

曖昧性をもつコロケーションは，大きく分けて３タイプあり，それぞれ，表１に示す要因によって生じていると考えられる。表１に示した例からも分かるが，曖昧性のタイプが(A)，(B)であれば，名詞の連体修飾句(表１では，「敵軍の」か「議長への」かという情報)によって，(C)であれば，従来どおりに格パターン（表１では，「シャワーデ」というデ格か，「額ニ」というニ格かという情報）によって，曖昧性の解消ができるということが分析結果から確認された。つまり，格パターン情報に加え，名詞の連体修飾句に関する情報も計算機に与えることが，コロケーションの曖昧性という問題解決に有効であるということである。

表１：コロケーションの曖昧性のタイプと要因

(A)名詞が多義で，述語は多義でない場合
[要因 A1]多義である名詞が同じ素性をもつため，素性に特有の同じ述語と結びつく。
（例）（敵軍の）攻撃[ACT]ガ始まる
（議長への）攻撃[ACT]ガ始まる
[要因 A2]述語の抽象度が高いため，同じ述語が多義である名詞と結びつく。
（例）（大豆の）収穫ガ多い
（会議の）収穫ガ多い
(B)名詞も，述語も多義である場合
[要因 B1]コロケーション自体が比喩的意味をもつため同じ名詞と述語の組合せがある。
（例）（口紅の）色ガ濃い
（あせりの）色ガ濃い
（敗戦の）色ガ濃い
(C)名詞は多義でなく，述語が多義の場合
[要因 C1]多義である述語が同じ名詞と結びつく。
（例）（シャワーデ）汗ヲ流す
（額ニ）汗ヲ流す
＊名詞と述語双方に多義性がない場合にはコロケーションの重複はない。

3.4 曖昧性解消に必要な辞書情報

しかしながら，曖昧性解消を十分に実現するには，名詞の連体修飾句情報を加えても，まだ十分ではなかった。

先に挙げた「骨が折れる」を例に，連体修飾句情報を使うことで，どのように曖昧性の解消が図れ，また，それ以外にどんな情報が必要とされるかを次に見てみる。図３に，ＩＰＡＬ名詞辞書の「骨が折れる」のコロケーションの記述を連体修飾句情報の記述と併せて示し，1)～5)に，実際のテキストから収集した用例を示す。

①高等動物において，体の支柱および枠組となる，カルシウムを多く含んだ硬いもの。
【ｺﾛｹｰｼｮﾝ】（手の／足の／…）骨[CON]が 折れる
【連体修飾句】
〈所有者〉{人，高等動物}老人の－，魚の－
〈部位〉{体の部分}頭の－，肩の－，足の－
②ものの芯となり，全体を支える役目を持つ部分
【ｺﾛｹｰｼｮﾝ】（傘の／…）骨[CON]が 折れる
【連体修飾句】
〈全体部分〉{樹状に支えるものがある物}傘の－，扇の－
③【慣用表現】骨が折れる（=苦労する）

図３：ＩＰＡＬ「骨」より「骨が折れる」の記述

1) 荒井さんは左足の【骨が折れ】，約三カ月のけが
2) 一色さんは頭の【骨が折れ】て重体。
3) ロシアを説得することが一番【骨が折れ】た
4) 根のない話を虚空に描き出すのには【骨が折れる】。
5)【骨が折れ】とるかも知れんな。

1)，2) は，「左足の」「頭の」と，体の部分を表す連体修飾句を伴っているので，①の連体修飾句の記述と照らし合わせることにより，①の意味だとすぐ分かる。一方，3)，4) は，①に示された体の部分を表す連体修飾句や，②に示された「傘」などの連体修飾句を伴わないことから①と②の可能性が低くなり，③であると考えられる。ところが，5) を，連体修飾句のない実例であることから，③と判断したとしたら，間違いであろう。

実は，3)～5)にある曖昧性は，さらに他の情報を参考にすることで解消が図れると考えられる。3) は，「一番」などの程度副詞である連用修飾句を伴い，4)は，ニ格をとり，かつ，そのニ格が「～には」という格表示をとっている。これらは③の慣用表現にしか現れ得ない統語的特徴であるということに着目できれば，3)，4) は③であるとはっきり判断できる。5)の場合，「骨が折れとる（ている）」というテイル形に着目できれば，慣用表現でないことは分かる。それは，「骨が折れる」の場合，③の慣用表現として用いられる時はテイル形をとりにくいと言えるからである。ただし，①か②であるかの曖昧性は，前後の文脈情報を参照しないと判断できない。

以上のことから，曖昧性解消には，単純な格パターン情報だけではなく，連体修飾句，連用修飾句，格表示といったこと全てに関する詳細な統語情報が必要であり，テイル形のような形態情報，さらには，文脈情報までが必要であることが予測される。なお，これは，コロケーションの曖昧性解消のためだけに限ったものではなく，語のレベルの曖昧性解消を十分に実現するためにも必要になってくる情報だと考える。

４．多義動詞の曖昧性解消過程の解明

３章で述べた通り，詳細な形態・統語・文脈情報というものが，語やコロケーションの曖昧性解消に必要になることは分かったが，情報の

内容は十分に明らかになっていない。そこで，現在，多義動詞を取り上げ，多義動詞の曖昧性解消過程の解明を目的とする研究を始めている。具体的には，二つの目的をたてている。第一に，多義動詞の用例を大量に収集し，そこから，曖昧性を解消し得るのに必要十分な，形態・統語・文脈情報の抽出を行うこと。第二に，抽出した情報をもとに，曖昧性解消過程の手続き化を試み，その手続き化の妥当性を計算機によって客観的に検証することである。

以下に，「見逃す」という動詞に関して，用例からの情報抽出と手続き化の例を示す。この「見逃す」は２つの意味をもつ語である。「見逃す」の実例を以下に示す。

1) 彼のサインを集団は【見のがし】ていたのである。
2) 若林は，原因を【見逃し】ていないし，
3) 高度情報化ということが含みもつ，こうした側面をも【みのがし】てはなりません。
4) どうかお【見逃がし】下さい
5) 日本の警官はあやまり方によって罪を【見逃し】てくれる。
6) たいていは一人四，五キロ，幸運な時は一〇キロぐらいのオーバーは【見逃し】てくれる。

以上の用例などを参照して，意味情報と，曖昧性を解消する手掛かりに成り得る，形態・統語情報を試みに記述してみると，図４のようになる（文脈情報の記述方法は，検討中）。

【意味情報】
①：見落とす。
②：見ないふりをして，とがめないでおく。
【形態情報】
①，②共通：見逃す・見逃がす・見のがす・みのがす
②：お見逃し下さい・見逃して下さい・見逃してくれる・見逃してやる・見逃そう
【統語情報】
①：＜人＞＋ガ＋＜可視物・事柄＞＋ヲ＋見逃す
＜可視物・事柄＞の例：サイン，原因，側面
②：＜人＞＋ガ＋＜罪・違反・誤り・悪事＞＋ヲ＋見逃す
＜罪・違反・誤り・悪事＞の例：罪，（速度）オーバー

図４：「見逃す」の意味・形態・統語情報

【手順】
［１］ 着目動詞の形態と形態情報②の記述とが一致すれば，その時点で②に確定。
［２］ 着目動詞のヲ格と統語情報のヲ格の記述を比較する。ヲ格の名詞句の意味範疇が②の方が狭いため，先に②であるかを判定し，判定できれば②に確定。それ以外で，①であるか判定できれば①に確定。判定が不確定，あるいは，着目動詞にヲ格の名詞句が無ければ，次の手順へ。
［３］ 文脈情報を利用して①か②に確定。

図５：「見逃す」の曖昧性解消の手順

また，抽出した情報を用いて解消を図る手順は，上の図５のようになる。

５．おわりに

多義という言語現象を大量の用例に基づいて分析することによって，多義語の意味理解に関し，国語学上重要な知見が得られると考えられる。また，今後，多義語の曖昧性解消過程を手続き化することで計算機モデルを構築し，曖昧性解消過程を検証する計画でいる。これにより，国語学的知見の客観性を吟味できるばかりでなく，結果として，計算機による多義性解消の精度向上にも寄与することが期待される。

参考文献

［ＥＤＲ93］日本電子化辞書研究所(1993)『ＥＤＲ電子化辞書仕様説明書』

[Fillmore75]Fillmore, C. (1975)『格文法の原理：言語の意味と構造』三省堂

［ＩＰＡ97］情報処理振興事業協会(1997)『CD-ROM版　計算機用日本語基本辞書ＩＰＡＬ－動詞・形容詞・名詞－』

［ＮＴＴ97］ＮＴＴコミュニケーション科学研究所（監修）(1997)『日本語語彙体系』岩波書店

［柏野・本多98a］柏野和佳子・本多啓(1998)「IPAL名詞辞書における多義構造の記述」『情報処理学会論文誌』39-9 pp. 2603-2612

［柏野・本多98b］柏野和佳子・本多啓（1998.12刊行予定）「多義構造を辞典に書く」『日本語学』17-14

［桑畑他98］桑畑（柏野）和佳子・橋本三奈子・青山文啓(1998)「IPAL名詞辞書による多義性解消のためのコロケーションの分析」『情報処理学会論文誌』39-6 pp. 1925-1934