JAIST Repository
https://dspace.jaist.ac.jp/

Title	HPSGを用いた古典ギリシア語文法の拡張
Author(s)	中嶋，健一郎
Citation	
Issue Date	2003-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1668
Rights	
Description	Supervisor:東条 敏，情報科学研究科，修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

HPSGを用いた古典ギリシア語文法の拡張

中嶋 健一郎 (110091)

北陸先端科学技術大学院大学 情報科学研究科

2003年2月14日

キーワード: HPSG, 構文解析, 古典ギリシア語, LiLFeS.

本論文には, 主辞駆動句構造文法 (HPSG) に基づく古典ギリシア語文法の拡張について述べられている. その目的は, 古典研究においてこれらの文献のテキスト情報だけでなく構文情報を扱えるようにすることである.

ギリシア古典と呼ばれる文献の多くは原本が既に失われており, 現在には写本だけが伝わっている. これらの写本は, これまで写字生などによって写本が繰り返されてきたという経緯の過程で, 写し間違いや個人の解釈を追加したりされてきたことなどにより, どこまでが原本なのか不明となってしまっているという問題がある. また, 現在には特定の著者によるものであると伝えられているものでさえ, 本当にその著者のみによるものか, もしくは複数の著者が存在していたのかということははっきりとしていない. こういった事柄に対して, 現代人の我々が著者の言い分を正確に汲み取るために行えることは, それらの文献から少しでも多くの情報を取得し, できるだけ古代ギリシャ人の考え方を客観的に研究することである. 古典ギリシア語の構文情報を取り出すことは, こうした研究の一助となることができると考える.

本研究で扱う, 古典ギリシア語では, 形容詞が名詞を格支配するとき, 付近ではなく離れた場所に出現できることや, 主語に限らず動詞, 目的語などであっても自明であるような要素ならば削除されるということ, 明らかに交差する係り受けが存在することなどという特徴が存在する.

Copyright © 2003 by Ken'ichiro Nakajima

本来構文情報を判断することは, 客観性が求められるにも関わらず, どうしても文法作成者の主観による判断を避けられない. この点においては, ユークリッドの『原論』は数学の証明の記述の集合であるというその性質上, 叙事詩などと比較し構造が論理的にすっきりとしている. そのため構文情報は比較的客観的に行うことが可能であると考えるため, 題材として適切であるといえる.

HPSG は多数の辞書記述を持ち, それぞれの辞書記述は豊富な情報からなる. そして, これらの語の情報を ID-schema, principle と呼ばれる少数の規則によって単一化を行い, 句を形成する. 古典ギリシア語はそれぞれの語に性·数·格など多数の属性を持つ. これらの属性を素性構造として扱い HPSG 上での単一化によって句を形成することにより, テキストの表層情報から構文情報を取り出すことができる.

HPSG 上で表現された他言語の文法において, 文法規則として規定されていることの多いものには, 主辞が補語を取って句を形成することを規定するものや句と句が修飾関係を形成することを規定するものなどがある. しかしながら, ある特定の品詞では語順の自由さのために, 句の配置が自由であるが, 他のものではそうではないなどの古典ギリシア語の特徴を考慮するとそのまま用いることはできず, このような現象を検討し, それを元に規則に修正を加えた.

これに加え, 既に出来上がっている句構造の内部に修飾関係を形成することを規定するスキーマや, 辞書上では主語をとる必要があるとされているものに対し, 主語を取らなくてもよいとする規則, 動詞を持たない文を生成する規則などを加えた.

更に, 名詞と形容詞の類似点と差異をどのように扱うか, 冠詞や接続詞の振舞いをどのように考えるべきかについて検討し, 語彙階層を再定義した. その結果, 本研究で作成した文法は, 合計 16 のスキーマと, 1354 語の辞書を持つ.

平均語長 16.9 語からなるユークリッドの『原論』7 巻 8 巻に対し, この文法を用いた結果 7 巻 673 文に対して 85.884%, 8 巻 481 文に対して 69.439%, 全体で 1154 文に対して 79.029%をのカヴァレッジを得ることができた. ま

た, 主語削除規則, 後方からの修飾規則を含まない文法, 後方からの修飾規則のみ含まない文法を用いた結果、7巻8巻について前者が69.710%, 及び49.576%, 後者が78.843%, 65.042%のカヴァレッジとなった. このことから, 本研究で定義したこれらの規則は古典ギリシア語文献を解析する上で有用であると言える. 解析時間については比較的長めの結果となった. この結果には, 表層情報のみから判断した場合における, 修飾関係のあいまい性が関係していると考える. 実装はLiLFeSの上で行ない, 構文解析機にはnaive CKY-parserを用いた.

本研究では, 文法記述から主観をできるだけ取り除くために, 意味情報をできるだけ用いず, 表層情報からの解析に終始した. そのため, あいまい性という点においては問題が残る. この文法に意味情報や, 文脈の論理性などから, 正しい係り受けや, 主語と補語の区別を判断できるようにすること, 辞書記述の追加, 改良をを加えることにより, 部分的な解析のあいまい性を減少させることができる. そうした結果として文全体のあいまい性を減少させることが今後の課題となる.