4K1-OS-2-9

句構造と依存構造について

On Phrase Structure and Dependency Structure

麻生 英樹[*1]
Hideki ASOH
[*1] 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology

Phrase structure analysis and Dependency structure analysis are two important approaches to analyze structures in natural language sentences and have been investigated for long period. In this short report, we investigate the relationship between them and address the possibility of integrating them.

1. はじめに

自然言語の文は文字あるいは音韻の 1 次元的な系列であるが，そこには文の意味や語の文法機能を反映した複雑な構造が内在している．書き言葉について考えると，文字列はまず単語という単位にまとめられる．さらにいくつかの単語がまとまりを作り，そうしたまとまり同志の間にも関係が存在する，といった形で，より複雑な構造が階層的に構築されていると考えられている [郡司 02]．こうした構造を分析，記述するために，古くから「依存構造」(dependency structure)と「句構造」(phrase structure)という二つのアプローチが提案され，研究されてきた．それぞれのアプローチに対応して，依存文法(たとえば [Tensniere 59, 児玉 87])と句構造文法(たとえば [Chomsky 56, Gazdar 85, Pollard 93]) という文法のクラスが存在し，それらの間の関係についても研究が行われている [Geifman 65, 児玉 87, Abney 94]．また，句構造文法における X バー理論 [Chomsky 70]，依存文法における語順の依存関係への取り込みなど，双方のアプローチを融合しようとする試みも多い．本稿では，句構造と依存構造という二つのアプローチを概観し，それらの間の関係と統合の可能性について検討する．

2. 依存構造と依存構造文法

依存構造は語や文節間の依存関係(修飾－被修飾関係，係り受け関係)に基づいて文に内在する構造を表現したもので，通常，単語や文節を節点とする木構造によって表現される．各節点には子節点から修飾を受ける被修飾語・文節が，その節点の子には親節点を修飾する(親節点に係る)修飾語・文節が置かれる．修飾される側を，この単位依存構造の主辞(ヘッド)，主要素などと呼び，修飾する側を従要素などと呼ぶ．文に対応する依存構造木全体は，こうした単位構造を接続して得られる．

「依存構造」が何を指すか，依存構造の単位は何か，などについては必ずしも合意があるわけではないが，たとえば [Robinson 70] は文の中の依存関係が満たす性質(公理)として以下のようなものを挙げている．

1) 1つの要素のみが独立している
2) 他の要素はすべて何らかの要素に直接依存する
3) どの要素も 2 つ以上の要素に直接依存しない
4) もし A が B に直接依存し，要素 C が線形順序をなす連鎖において両者の間に位置する場合，C は A または B または A と B の間に位置する他の要素に直接依存する．

ただし，このうちの 3) と 4) については依存関係の公理に含めるか否かについて意見が分かれているようである．また，児玉は，主要素と従要素を特徴づける現象として下記を挙げている [児玉 87]

1) 要素の A と B が結合した場合，範疇において要素群 AB の特徴的役割をはたす要素が主要素，他方が従要素である．
2) 語順において従要素の位置は主要素の位置に応じて決定され，その逆ではない．
3) 主要素は共起する従要素の数と種類を決定する．
4) 従要素の意味は主要素の意味構造に含まれるものであり，その逆ではない．
5) 主要素が従要素の屈折形態を決定する．
6) 主要素を残して従要素を省略することができるが，その逆は不可能である．

依存構造に基づいて，自然言語の文法を定義する試みは依存文法と呼ばれる．児島はそうした試みを以下のような様々な観点から分類している [児玉 87]．

1) 節点に語彙項目を置くものと，語彙範疇を置くもの
2) 文の主要素に modality を取るものと，動詞を取るもの
3) 変形規則を持つものと持たないもの
4) 語順を依存構造に含めるものと，依存構造以外の情報も含めて総合的に語順を決めるもの．
5) 格関係の扱いなど依存関係自体の細かい違い．
6) 意味情報と依存関係をどのように関連づけるかの違い．

日本語やヨーロッパの言語など，明示的な格標識を持ち語順が比較的自由とされている言語においては自然言語処理の構文解析として，句構造ではなく依存構造を得ることがよく行われている．日本語の依存構造解析器としては，たとえば，CaboCha [工藤 02]，KNP [河原 07] などがよく知られている．

3. 句構造と句構造文法

句構造は隣接する語の間の関係に基づいて文の構造を表現したもので，通常，句構造文法の生成規則(句構造規則 phrase structure rule)によって定義される．句構造は語および文法的カテゴリ(非終端記号)を節点とする木構造で表される．句構造規則群を中心とする文法を句構造文法(phrase structure grammar)と呼ぶ．文から句構造規則に矛盾しない句構造木を生成する操作を構文解析(parsing)と呼ぶ．このように，句構造とは，文中の語を，意味・機能的まとまりおよび隣接構造の二つの側面から眺めて得られる構造である．英語など語順によって語の文法的および意味的役割が規定される言語は，句構造文法による構文解析が行われることが多い．

自然言語は文脈自由句構造文法の範囲では規定されないと考えられているため，文脈自由句構造文法を拡張して自然言語の文法を構築する研究が盛んに行われた．Chomsky は句構造規則で作られた構造に変形操作を加えて表層構造が作られるとした[Chomsky 56]．Gazdar らは統語範疇に素性構造を導入することで文法を拡張した[Gazdar 85]．この方向はその後，

連絡先：麻生英樹，産業技術総合研究所 知能システム研究部門，〒305-8568 つくば市梅園 1－1－1 中央第二

Pollard らによって，主辞駆動句構造文法(HPSG)へと推し進められた[Pollard 93, Sag 03]．また，古典的な範疇文法は文脈自由句構造文法の一種と考えられるが，それを拡張した組み合わせ範疇文法 [Steedman 96, 戸次 10] も提案されている．

4. 句構造と依存構造の関係と統合

句構造と依存構造は通常どちらも統語的構造として考えられており，句構造文法と依存文法の間の関係，句構造と依存構造の間の相互変換などが研究されてきている．たとえば，Geifman は，語順も考慮した依存構造に基づく依存文法が文脈自由句構造文法のサブセットであることを示した [Geifman 65]．ただし，この結果については批判もある．たとえば Abney は，主辞性を含めた句構造と依存構造の対応関係について検討し，依存構造文法が句構造文法のサブセットとは言えないと主張している [Abney 94]．また，Jarvinen らも，Geifman の定式化による依存文法と Tesniere によるものとの乖離を指摘している [Jarvinen 98]．

表面的には，句構造と依存構造の違いは，中間構造の有無である．句構造規則はたとえば名詞句（NP），動詞句（VP），前置詞句（PP）等の中間構造を生成するが，依存構造は基本的に単語間の依存関係であり，そうした中間構造は導入されない．しかし，依存構造もまた木構造の形に表現されるのであり，ある節点以下の語がひとつのまとまりとしての中間構造を作り，それ全体として上位の節点の語を修飾していると見なすこともできる．このように考えると，中間構造の有無は，必ずしも句構造と依存構造の本質的な違いではないとも言える．

句構造と依存構造のもうひとつの違いは，語の隣接関係，語順の考慮である．句構造規則では，ある句構造はいくつかの隣接する要素構造から構成される．一方，依存構造においては，依存関係にある語は必ずしも隣接しているとは限らないし，語順もまた規定されないとすることが一般的である．句構造規則に含まれる隣接関係は，交差的な依存や遠隔依存などが頻出する言語の文法を記述する際に問題になることがある．

句構造と依存構造を統合する方向の研究も行われてきた．その中で最も重要なものは句構造文法への「主辞性」の導入であろう．一般的な句構造規則には主辞性という概念は無い．これに対して，[Chomsky 70] 以降に変形生成文法理論に導入された X バー理論は，様々な統語範疇に関する句構造規則を抽象的なレベルでまとめるとともに，句構造規則に主辞性の概念を導入したものと考えることができる．さらに，Gazdar らの GPSG では，句構造規則の右辺に，主辞を表す H という記号を導入することによって，より直接的に句構造規則に主辞の概念を導入した [Gazdar 85]．その考え方は HPSG にも継承され，発展している [Pollard 93, Sag 03]．

句構造文法や範疇文法で導入されている語彙化すなわち，語彙に多くの情報を持たせて，生成規則を簡素にするという方向も，語や語と語の関係を重視する依存構造的な考え方を句構造文法に持ち込む試みとも言えるだろう．また，Jackendoff らのように，制約充足の考え方を用いて，構造と語順をできるだけ分けて規定しようという試みもある [Culicover 05]．

5. おわりに

自然言語の文の構造を捉えるための，句構造と依存構造という二つのアプローチを概観し，それらの間の関係について検討した．主辞性を導入し語彙化の方向を強めた句構造文法によって，二つのアプローチはほぼ統合されたようにも思われるが，その一方で，依然として語順の扱いなどでは意見が分かれている面もある．また，統語と意味の関係についても，戸次による日本語 CCG のような有望な方向が示されているものの，確立された結論があるわけではない．依存構造は語と語の修飾関係を基本としているため，語の意味との親和性が高い．実際，依存関係をベースとするさまざまな意味表現が提案されている．古くは，Schank による概念依存関係があるが [Schank 75]，その他にも，たとえば，語彙意味論における型つきの関数による意味表現 [Culicover 06]や，形式意味論における述語論理式による意味表現がある[戸次 10]．それらにおいて，関数とその引数の関係や，述語とその項の関係は，主辞とそれを修飾する語との依存関係に他ならない．その一方で，それらの意味表現においては，関数や述語が別の関数や述語を引数に取るような高階表現が可能であり，たとえば関数の型は句構造的な中間構造に対応していると考えることもできる．二つのアプローチを統合した言語普遍的な枠組みに向けて，さらなる研究が必要である．

謝辞：本研究の一部は科研費 23240043 の助成を受けた．

参考文献

[郡司 02] 郡司孝雄: 単語と文の構造, 岩波書店, 2002.

[Tensniere 59] L. Tensniere: Elements de Syntaxe Structurale, Kliensieck, 1959.（小泉（監訳），構造統語論要説，研究社，2007.）

[児玉 87] 児玉徳美: 依存文法の研究，研究社出版, 1987.

[Chomsky 56] N. Chomsky: Three models for description of language, IRE Trans. Information Theory, IT-2, 113-124, 1956.

[Gazdar 85] G. Gatzdar, E. Klein, G. Pullum, and I. A. Sag: Generalized Phrase Structure Grammar, Basil Blackwell Publisher, 1985.

[Pollard 93] C. Pollard and I. A. Sag: Head-Driven Phrase Structure Grammar, CSLI Publications, 1993.

[Geifman 65] H. Gaifman: Dependency systems and phrase-structure systems, Information and Control, 8, 304-337, 1965.

[Abney 94] S. Abney: Dependency grammars and context-free grammars, University of Tuebingen, Draft, 1994.

[Chomsky 70] N. Chomsky: Remarks on Nominalization, Jacobs, R. A. and Rosenbaum, P. S. (eds.), Reading in English Transformational Grammar, 184-221, Ginn & Co., 1970.

[Robinson 70] J. J. Robinson: Dependency structures and transformational rules, Language, 46, 259-285, 1970.

[工藤 02] 工藤, 松本: チャンキングの段階適用による日本語係り受け解析, 情報処理学会論文誌, 43, 6, 1834-1842, 2002.

[河原 07] 河原, 黒橋: JUMAN/KNP を用いた形態素・構文・格解析, 京都大学学術情報メディアセンター, 2007.

[Sag 03] I. A. Sag, T. Wasow, E. M. Bender: Syntactic Theory, A Formal Introduction (2nd Ed.), CSLI Publications, 2003.

[Steedman 96] M. Steedman: Surface Structure and Interpretation, The MIT press, 1996.

[戸次 10] 戸次大介: 日本語文法の形式理論, くろしお出版, 2010.

[Jarvinen 98] T. Jarvinen and P. Tapanainen: Towards an implementable dependency grammar, arXiv:cmp-lg/9809001v2, 1998.

[Culicover 05] P. W. Culicover and R. Jackendoff: Simpler Syntax, Oxford University Press, 2005.

[Schank 72] R. Schank: Conceptual dependency, a theory on natural language understanding, Cognitive Psychology, 3, 522-631, 1972.