2009 年冬の LA シンポジウム [S9]

語彙作用文法による自然言語の構文解析

著者 島村 瑛* 著者 笠井 琢美†

1 まえがき

英語などの自然言語には曖昧性があり、1つの文が複数の構造を持ち得る。特に、現在の生成文法の理論では修飾子がどの語句に係るかを正確に述べることはできない。そのため、構文解析の際に全ての解析木を得ようとすると解析木の爆発が起こり得るという問題がある[4]。

(1) I saw a girl with a telescope.
（**a**）**(I (saw (a girl with a telescope.)$_{NP}$)$_{VP}$)$_S$**
私は望遠鏡を持っている少女を見た。
（**b**）**(I (saw (a girl)$_{NP}$ with a telescope.)$_{VP.}$)$_S$**
私は少女を望遠鏡で見た。

(1)はwith句が 名詞 **girl** を修飾するか動詞 **saw** を修飾するかで（**a**）と（**b**）の2通りの構造に解析できる。さらに、with句の後ろに前置詞句が現れると、その前置詞句の修飾先は3通りとなる可能性があり、計6通りとなる。このように、修飾子が係り得る全ての場合に対して解析木を求めようとすると、修飾子の数に対して出力される解析木の数が指数関数的に増加してしまい、破綻する。

本論文では文の代表となる構造(核構造) だけを生成するため、語彙作用文法と呼ばれる新しい文法モデルを提案する。語彙作用文法では、素性値を用いて各修飾子の修飾先を1通りに制限する。語彙作用文法によって生成される構造を核構造と呼ぶ。使用者は核構造から文を派生させて目標の構造を得ることができる。英語では多くの単語が複数の“型”を持つ。例えば、動詞と名詞の両方の用法を持つ単語は多い。英語の各単語に添え字を付け、それぞれが1つの型を持ったようにした言語を添え字付加言語と呼ぶ。このようにして得られる英語の添え字付加言語は“左再帰自由”と呼ばれる特別な性質を有する。本論文では左再帰自由な添え字付加言語は決定性言語であることを証明する。

*電気通信大学大学院情報工学専攻
†電気通信大学情報工学科

2 語彙作用文法の定義

2.1 語彙作用文法

語彙作用文法の形式的な定義を示す。

[定義 1] 語彙作用文法とは、5 項組 $G = (\Gamma, \Sigma, \Delta, CAT, S)$ のことである。ここで、Γ, Σ, Δ は空でない有限集合である。Γの元を原始型、Σの元を終端記号、Δの元を素性値という。S は文標識と呼ばれる特別な原始型である。G は演算子と呼ばれる特別な記号/. \, →,←, & を常に含む。

G の型を帰納的に定義する。

(1) 基本型αとΔの部分集合 K に対し、$\alpha\&K$ は型である。

基本型は原始型か次で定義される還元型または修飾型のことである。

(2) 基本型α, 原始型βに対し、α/β, $\beta\backslash\alpha$は還元型である。

(3) 基本型α, 還元型βに対し、α/β, $\beta\backslash\alpha$は還元型である。

修飾型は次で定義される。

(4) 基本型αに対し、$\rightarrow\alpha$, $\leftarrow\alpha$は修飾型である。

G の型$\alpha\&K$において $K = \phi$ である場合αと書くことにする。

また、還元型$(...((\alpha/\beta_0)/...)/\beta_n$における$\alpha$を右最終形と呼び、修飾型$\rightarrow(....\rightarrow(\rightarrow(\alpha))...)$における$\alpha$を左修飾最終形と呼ぶ。

Gの基本型全体からなる集合と型全体からなる集合を T^G、Ω^G で表す。ただし、Gが明らかなとき、T^Gを T、Ω^Gを Ωで表す。CATはΣから型の有限集合の族への関数である。

これ以降、G は語彙作用文法 $G = (\Gamma, \Sigma, \Delta, CAT, S)$ を表すものとする。

$\Omega^* \times \Sigma^*$の元をGの様相という。$(y; w)$を様相とするとき、wを入力列、yを還元列と呼ぶ。

[定義 2] A_1, A_2を Ω^*の元、α, β, γを Tの元、K_1, K_2 をΔの部分集合、aをΣの元、wをΣ上の文字列とする。

Tから型の有限集合の族への関数g_{m}を修飾認可関数、Tから型の有限集合の族への関数g_{r}を還元認可関数、$\Omega\times\Omega$から素性値の有限集合の族への関数fを素性値伝搬関数という。
Gの様相上の関係$\vdash_G$は次で定義される。

（読込み規則）
$\alpha\&K_1\in CAT(\mathrm{a})$ならば
$(A_1;\mathrm{a}\ w)\vdash_G(A_1,\ \alpha\&K_1;w)$.

（右還元規則）
$\gamma\&K_2\in g_{\mathrm{r}}(\beta)$ならば
$(A_1,\alpha/\beta\&K_1,\gamma\&K_2,A_2;w)\vdash_G(A_1,\ \alpha\&K_3,A_2;w)$
ここで、$K_3=f(\alpha/\beta\&K_1,\ \gamma\&K_2)$.

（左還元規則）
$\gamma\&K_1\in g_{\mathrm{r}}(\beta)$ならば
$(A_1,\ \gamma\&K_1,\ \beta\backslash\alpha\&\mathrm{K2},A_2;w)\vdash_G(A_1,\ \alpha\&K_3,A_2;w)$
ここで、$K_3=f(\gamma\&K_1,\ \beta\backslash\alpha\&\mathrm{K2})$.

（右修飾規則）
$\beta\&K_1\in g_{\mathrm{m}}(\leftarrow\beta)$ならば
$(A_1,\ \beta\&K_1,\leftarrow\beta\&K_2,A_2;w)\vdash_G(A_1,\ \beta\&K_3,A_2;w)$
ここで、$K_3=f(\beta\&K_1,\leftarrow\beta\&K_2)$.

（左修飾規則）
$\beta\&K_2\in g_{\mathrm{m}}(\rightarrow\beta)$ならば
$(A_1,\rightarrow\beta\&K_1,\ \beta\&K_2,\mathrm{A2};w)\vdash_G(A_1,\ \beta\&K_3,A_2;w)$
ここで、$K_3=f(\rightarrow\beta\&K_1,\ \beta\&K_2)$.

$\vdash_G$の反射推移閉包を$\vdash_G^*$と表す。Gの還元とは、$\beta_0\vdash_G\beta_1\vdash_G\cdots\vdash_G\beta_n$で実現される様相の列$\beta_0,\beta_1,\cdots,\beta_n$である。$G$が明らかなとき、$\vdash_G$と$\vdash_G^*$を単に$\vdash$と$\vdash^*$と書く。$G$の受理する言語$L(G)$を$\Sigma$上の言語$\{w\in\Sigma^*\mid(\varepsilon;w)\vdash^*(S\&K;\varepsilon)\}$と定義する。$\Sigma$上の言語$L$が語彙作用文法によって受理されるとき、$L$を語彙作用言語と呼ぶ。

以下で、定義 2 の各規則について説明する。読込み規則において、関数CATは、各終端記号に有限個の許される型を定める。$(A_1;\mathrm{a}\ w)\vdash_G(A_1,\alpha\&K_1;w)$は、入力列の左端からaを取り除き、aに許された型$\alpha\&K_1$を還元列の右端に加える操作を表す。右還元規則において、還元認可関数g_{r}は還元型に対となることのできる有限個の型を定める。$(A_1,\alpha/\beta\&K_1,\gamma\&K_2,A_2;w)\vdash_G(A_1,\alpha\&K_3,A_2;w)$は還元列中に現れた型$\alpha/\beta\&K_1$と型$\gamma\&K_2$を対とし、$\alpha/\beta\&K_1,\ \gamma\&K_2$を型$\alpha\&K_3$に書き換える操作を表す。素性値伝搬関数$f$は還元が行われる型の対から還元後の型の素性値の集合を定める。修飾の制限に関わる素性値は修飾子素性と呼ぶ。修飾子素性には強さがあり、同じ型の修飾に作用する素性値はより修飾を強く制限する素性値が伝搬される。例えば、英語において人称代名詞は普通の名詞よりも修飾されにくく、人称代名詞を修飾できる名詞修飾子と修飾できない名詞修飾子がある。そのため人称代名詞からは普通の名詞よりも弱い素性値が伝搬され、より強い修飾子素性に置き換えられる可能性がある。同様に、左還元規則において、$(A_1,\gamma\&K_1,\beta\backslash\alpha\&\mathrm{K2},A_2;w)\vdash_G(A_1,\alpha\&K_3,A_2;w)$は還元列中に現れた型$\beta\backslash\alpha\&\mathrm{K2}$と型$\gamma\&K_1$を対とし、$\gamma\&K_1,\ \beta\backslash\alpha\&K_2$を型$\alpha\&K_3$に書き換える操作を表す。右修飾規則において、還元認可関数g_{m}は修飾型に対となることのできる有限個の型を定める。$(A_1,\beta\&K_1,\leftarrow\beta\&K_2,A_2;w)\vdash_G(A_1,\beta\&K_3,A_2;w)$は還元列中に現れた型$\beta\&K_1$と型$\leftarrow\beta\&K_2$を対とし、$\beta\&K_1,\ \ \leftarrow\beta\&K_2$を型$\beta\&K_3$に書き換える操作を表す。同様に、左修飾規則において、$(A_1,\rightarrow\beta\&K_1,\beta\&K_2,\mathrm{A2};w)\vdash_G(A_1,\beta\&K_3,A_2;w)$は還元列中に現れた型$\beta\&K_1$と型$\beta\&K_2$を対とし、$\rightarrow\beta\&K_1,\ \beta\&K_2$を型$\beta\&K_3$に書き換える操作を表す。

2.2 修飾の制限

以下に**(1)**における素性値による修飾の制限の例を示す。

saw に他動詞を意味する型 VP/NP, **a girl** に名詞句を意味する型 NP, **with a telescope** に VP と NP を右から修飾できる型 ←NVP が与えられているとする。（**b**）の構造を得ようとする場合、まず **saw** と **a girl** に与えられた VP/NP と NP に右還元規則が適用される。その際に素性値の伝搬が行われ、**saw a girl** の型は VP&{<NP 済>} となる。<NP 済>は **girl** より後に現れた名詞修飾子は **girl** が引き受けたため、名詞修飾子では修飾されないことを意味する素性値である。そのため、名詞修飾子でもある ←NVP の <NP 済>, <VP 済>などを素性値に持つ型への修飾は認可されない。よって VP&{<NP 済>} と ←NVP では右修飾規則が適用されず、最終的に解析は失敗する。一方で、（**a**）の構造を得ようとする場合にはまず NP と ←NVP の還元が試される。上記の素性値を持たない NP への ←NVP の修飾は認可されるため、NP と←NVP に右修飾規則が適用される。そのまま解析は進み、（**a**）の構造が**(1)**の核構造として生成される。**図 1** に**(1)**の核構造を示す。

3 添え字付加言語の決定性について

3.1 左再帰自由

左還元型$\alpha\backslash\beta$のαを左還元子、βを左被還元子と呼ぶ。

[定義3] Gが受理する言語が左再帰自由であるのはGの基本型が以下の条件を満たすときである。

(1) αがある型の左還元子として現れたら、任意の型の左被還元子として現れることができない。
(2) 修飾型は左被還元子として現れることができない。

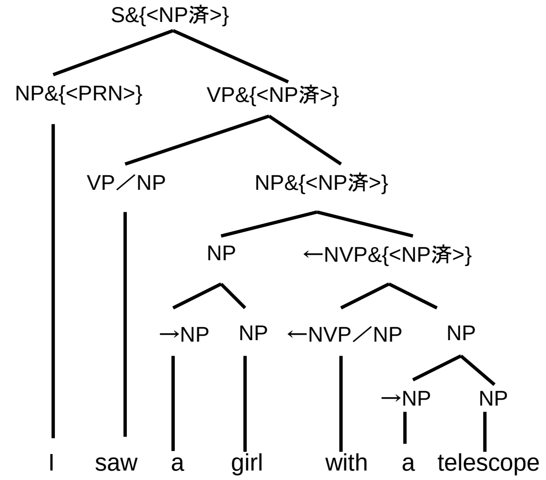

図 1. **I saw a girl with a telescope.**の核構造

3.2 添え字付加言語

英語では1つの単語に複数の品詞が与えられていることが多い。一般に自然言語では、1つの終端記号には複数の型が与えられる。例えば、終端記号**walk**には名詞句を表す型NP, 自動詞を表す型VP,他動詞を表す型VP/NPが与えられる。つまり、*CAT*(walk) = {NP, VP, VP/NP} となる。ここで、**walk**に添え字を付加してそれぞれに型を1つだけ与える。このとき、*CAT*はΣからΩへの関数として扱うことが出来る。つまり、$CAT(\mathrm{walk}_1)$ = NP, $CAT(\mathrm{walk}_2)$ = VP, $CAT(\mathrm{walk}_3)$ = VP/NP となる。このように、全ての終端記号に型を1つだけ与えた語彙作用言語を添え字付加言語と呼ぶ。

[事実1]

英語の添え字付加言語は左再帰自由である。

生成文法において等位構造を生成する英語の等位接続詞はX\X/X の型で表されるが、語彙作用文法では←X/X という修飾型で実現される。そのため、英語の語彙作用文法では、左還元型は左に現れた主語と還元されて節になることを表す型NP\Sしか存在しない。よって、英語の添え字付加言語は左再帰自由となる。

古典的英文法の主要8品詞[5]に対応する語彙作用文法の型の例を**表1**に示す。

表1. 英語の品詞と語彙作用文法の型の対応

英語の品詞	語彙作用文法の型
名詞	NP
代名詞	NP&{<PRN>}
形容詞	ADJ, →NP, ←NP
動詞	NP\S (VP), VP/NP, etc
副詞	←VP, →VP
前置詞	←NVP/NP
接続詞	←VP/S, NP/S, ←X/X
間投詞	←VP, ←S, →S

[定理1] 左再帰自由な添え字付加言語は決定性言語である。

（証明）語彙作用文法が受理する言語を左再帰自由な添え字付加言語とすると、定義 2 で示した様相上の関係$\vdash_G$を以下のように拡張することができる。

[定義 4] A_1をΩ*の元、$\alpha, \beta, \gamma, \chi$を$T$の元、$K_1$, K_2, K_3をΔの部分集合、aをΣの元、*w*をΣ上の文字列とする。Tから型の有限集合の族への関数g_{m}を修飾認可関数、Tから型の有限集合の族への関数g_{r}を還元認可関数、$\Omega \times \Omega$から素性値の有限集合の族への関数*f*を素性値伝搬関数、TからTへ関数*rrf*を右最終形関数、TからTへ関数*lmf*を左修飾最終形関数という。

拡張された様相上の関係$\vdash$は次で定義される。

（拡張読込み規則）

下記の規則が適用できない ∧ $\alpha \& K_1 = CAT(\mathrm{a})$ ならば

$$(A_1; \mathrm{a}\, w) \vdash_G (A_1,\ \alpha \& K_1; w).$$

（拡張右還元規則）

$\gamma \& K_2 \notin g_{\mathrm{m}}(\chi) \wedge \gamma \& K_2 \notin g_{\mathrm{m}}(rrf(\chi)) \wedge$
$\gamma \& K_2 \notin g_{\mathrm{m}}(lmf(\chi)) \wedge \gamma \& K_2 \in g_{\mathrm{r}}(\beta)$ならば

$$(A_1,\ \alpha/\beta \& K_1,\ \gamma \& K_2,\ \chi \& K_3; w)$$
$$\vdash_G (A_1,\ \alpha \& K_4,\ \chi \& K_3; w)$$

ここで、$K_4 = f(\alpha/\beta \& K_1,\ \gamma \& K_2)$.

（拡張左還元規則）

$\beta \backslash \alpha \& K_2 \notin g_{\mathrm{m}}(\chi) \wedge \beta \backslash \alpha \& K_2 \notin g_{\mathrm{m}}(rrf(\chi)) \wedge$
$\gamma \& K_2 \notin g_{\mathrm{m}}(lmf(\chi)) \wedge \gamma \& K_1 \in g_{\mathrm{r}}(\beta)$ならば

$$(A_1,\ \gamma \& K_1,\ \beta \backslash \alpha \& K_2,\ \chi \& K_3; w)$$
$$\vdash_G (A_1,\ \alpha \& K_4,\ \chi \& K_3; w)$$

ここで、$K_4 = f(\gamma \& K_1,\ \beta \backslash \alpha \& K_2)$.

（拡張右修飾規則）

$\leftarrow\beta \& K_2 \notin g_{\mathrm{m}}(\chi) \wedge \leftarrow\beta \& K_2 \notin g_{\mathrm{m}}(rrf(\chi)) \wedge$
$\gamma \& K_2 \notin g_{\mathrm{m}}(lmf(\chi)) \wedge \beta \& K_1 \in g_{\mathrm{m}}(\leftarrow\beta)$ならば

$$(A_1,\ \beta \& K_1, \leftarrow\beta \& K_2,\ \chi \& K_3; w)$$
$$\vdash_G (A_1,\ \beta \& K_4,\ \chi \& K_3; w))$$

ここで、$K_4 = f(\beta \& K_1, \leftarrow\beta \& K_2)$.

（拡張左修飾規則）

$\beta \& K_2 \in g_{\mathrm{m}}(\rightarrow\beta)$ならば

$$(A_1, \rightarrow\beta \& K_1,\ \beta \& K_2,\ \chi \& K_3; w)$$
$$\vdash_G (A_1,\ \beta \& K_4,\ \chi \& K_3; w)$$

ここで、$K_4 = f(\rightarrow\beta \& K_1,\ \beta \& K_2)$.

G の初期様相は $(\varepsilon; w, \$)$ で、受理様相は $(S\&K, \$; \varepsilon)$である。入力列の$\$$は入力の終わりを意味する終端記号、還元列の $\$$ は還元されない型である。また、右最終形関数 *rrf* は右還元型の右最終形を求め、*lmf* は左修飾型の左修飾最終形を求める関数である。

拡張された様相上の関係では、還元列をスタックとして考えると、還元は必ずトップより1段深

いところで行われる。これは、入力を1つ先読みしていることを意味している。拡張された右還元規則、左還元規則、右修飾規則では、規則を適用する条件として、還元列の右端の型がその左の型を右から修飾できないことを加えている。還元列の右端の型がその左にある型を修飾できるかどうかは、定義3の(2)より、その型の右最終形や左修飾最終形を見ることで判断できる。拡張右還元規則について説明すると、還元列中に α/β, β が現れた場合、定義3の(1), (2)より、β 以降の還元によって再び β が現れることはないので、β が後ろから修飾されない限り、α/β と β に右還元規則を適用できる。つまり1語先読みすれば右還元規則を適用するかを決定できる。拡張読み込規則において、添え字付加言語は、その定義から各終端記号に許される型は1つだけとなり、決定的に終端記号に与えられる型を定めることができる。拡張された関係は、還元列に加えられた型が以下の順に処理されることを意味している。まず拡張左還元規則によって左にある型からの修飾が処理される。その次に拡張規則によって右にある型からの修飾が処理される。その次に、左にある型ある型とで拡張右還元規則もしくは拡張左還元規則が適用できるか試される。規則が適用できない場合には、拡張読み込規則によって、次の入力を読み込む。

このように、拡張された様相上の関係について、任意の様相に適用される規則はただ一つに定まる。よって、左再帰自由な添え字付加言語は決定性言語である。

事実 1、定理 1 から英語の添え字付加言語は決定性言語であることが言える。

[例] $\mathbf{I_1}$ **saw**$_1$ **a**$_1$ **girl**$_1$ **with**$_1$ **a**$_1$ **telescope**$_1$.という文について考える。語彙作用文法が、 I_1, saw_1, a_1, $girl_1$, $with_1$, $telescope_1$ の 6 つを終端記号としてもち、S とNPの 2 つを原始型としてもち、<NP済>を素性値としてもつとする。ここで、$CAT(I_1)$ = NP, CAT(saw1) = (NP\S)/NP, CAT(a1) = →NP, CAT(girl1) = NP,$CAT(with_1)$ = ←NVP/NP, $CAT(telescope_1)$ = NP とする。この文の還元は次のように行われる。

(ε ; I_1, saw_1, a_1, $girl_1$, $with_1$, a_1, $telescope_1$, $)
⊢ (NP; saw_1, a_1, $girl_1$, $with_1$, a_1, $telescope_1$, $)
⊢ (NP, (NP\S)/NP; a_1, $girl_1$, $with_1$, a_1, $telescope_1$, $)
⊢ (NP, (NP\S)/NP, →NP; $girl_1$, $with_1$, a_1, $girl_1$, $)
⊢ (NP, (NP\S)/NP, →NP, NP; with1, a1, girl1, $)
⊢ (NP, (NP\S)/NP, →NP, NP, ←NVP/NP; a_1, $girl_1$, $)
⊢ (NP, (NP\S))/NP, NP, ←NVP/NP; →NP, $girl_1$, $)
⊢ (NP, (NP\S))/NP, NP, ←NVP/NP, →NP; $girl_1$, $)
⊢ (NP, (NP\S))/NP, NP, ←NVP/NP; →NP, NP; $)
⊢ (NP, (NP\S))/NP, NP, ←NVP/NP; →NP, NP, $; ε)
⊢ (NP, (NP\S)/NP), NP, ←NVP/NP, NP, $; ε)
⊢ (NP, (NP\S)/NP, NP, ←NVP/NP, NP, $; ε)
⊢ (NP, (NP\S)/NP, NP, ←NVP&{<NP 済>}, $; ε)
⊢ (NP, (NP\S)/NP, NP&{<NP 済>}, $; ε)
⊢ (NP, NP\S&{<NP 済>}, $; ε)
⊢ (S&{<NP 済>}, $; ε)

4 英和翻訳システム

4.1 eInG による木の操作

当研究室では英和翻訳システム「**eInG**」を開発している。自然言語の形式化には、文脈自由文法よりもやや強力な文法モデルが必要であるとされている[3]。**eInG**では、tree adjoining grammars と等価な転送スタック付きプッシュダウンオートマトン[1]と呼ばれるオートマトンモデルで構文解析を行っている。語彙作用文法の生成能力は文脈自由文法と等しいが、語彙作用文法もこのモデルに自然に拡張することができ、**eInG**上にて実現されている。

eInG にて**(1)**を入力すると、図**2**に示す**(1)** の(a)に対応する解析木とそれに対応する訳「望遠鏡を持った少女を見ました。」のみが出力され、(b)に対応する解析木と訳は出力されない。

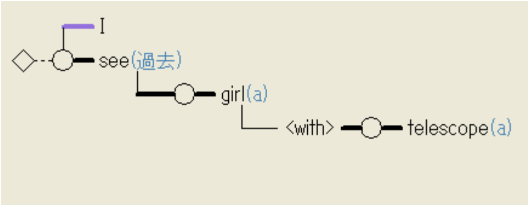

図 2. eInG による(1) の解析木

我々の解析木では単語自身が木の節点となっている。一番左にある◇が解析木の根となっており、生成文法の CP に対応している。図 2 の解析木において girl から with 句へ枝が伸びていることで、with 句が girl に係っていることを表している。画面上の with をダブルクリックして現れる図 3 のメニュー項目の内、上昇を選択すると、図 4 が示す**(1)** の(b)に対応する解析木へと変化し、対応する訳も「望遠鏡で少女を見ました。」となる。上昇は選択された修飾子の節点の親を、現在の親を起点にしてそこから根へ辿って行く中で初めて見つけた修飾可能な節点に変更する操作を表している。また、上昇とは逆の操作を表す下降がある。**eInG** の使用者はこれらの操作を用いることで、構文解析において制限されていた修飾先に修飾子をかけ直すことができ、目標の構造が得られる。

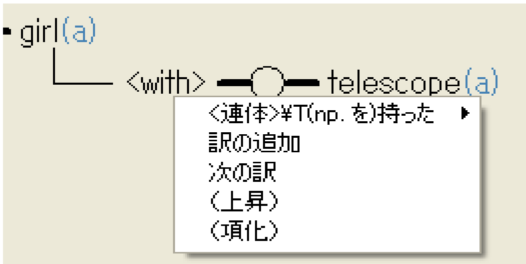

図 3. 解析木の節点の操作項目

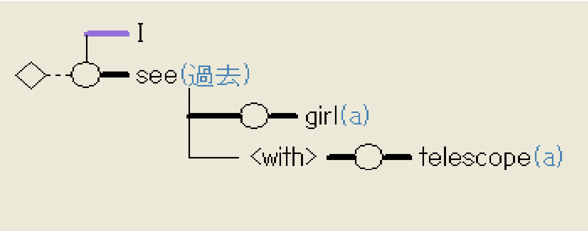

図 4. with 句の上昇操作後の解析木

(2) he asked his doctor for advice.
彼は医者に助言を求めました。

(2)の英文において、for句は ask と結びつきが強く、ask にとってほぼ必須の要素となっている。このような要素を項という。 for句は ask の項であるという情報をあらかじめ辞書に登録しておくことができる。その情報から構文解析後の再解析においてfor句の修飾先を doctor から ask に変更し、初めから for句が ask に係っている図5の木を出力することができる。このように、核構造をよりもっともらしい構造にするために、修飾子の修飾先を決定するための情報を収集することも今後の課題となっている。

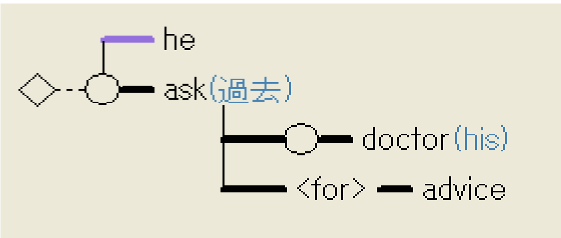

図5 eIngによる(2)の解析木

4.2 素性値による型の削減

実際の英語は添え字付加言語ではない。よって各単語には複数の型が与えられることになる。そのため、修飾子の曖昧性による解析木の増加が抑えられるようにはなったが、(3)のように各単語の型の組み合わせによっても文が複数の構造を持つことがある。

(3) Flying planes can be dangerous.
(a) 飛んでいる飛行機は危険なことがある。
(b) 飛行機を飛ばすことは危険なことがある。

fly に自動詞の型 VP が与えられた場合には flying は →NP として planes を修飾する。この構造に対応する訳は(a)となる。一方で、fly に他動詞の型 VP/NP が与えられた場合の flying は planes を目的語に取って、NPとして文の主語となる。この構造に対応する訳は(b)となる。

型の組み合わせによる解析木の増加を抑えるため、また構文解析を効率良く行うために、一つの単語に与えられる型の数を抑えることが今後の課題となっている。そのために、1つの型で複数の型の役割を果たす新しい型を考える。そのうちの1つが時の副詞名詞の型である。yesterday などの時を表す名詞は文中で名詞として働く場合と副詞として働く場合がある。このような名詞を副詞名詞と呼ぶ。語彙作用文法では、型 NP の素性値に〈時〉を加えた型 NP&{〈時〉}で時の副詞名詞を表す。以下に時の副詞名詞を用いた例文を示す。

(4) Yesterday I visited the zoo.
昨日、動物園に行きました。
(5) I met you yesterday.
私は、昨日、あなたに会いました。
(6) Yesterday was Sunday.
昨日は日曜日でした。

(4), (5) では yesterday は副詞として働き、(6) では名詞として働いている。型 X/NP の次に副詞名詞の型 NP&{〈時〉} が現れた場合や副詞名詞の型が文頭に現れ、その次に動詞が現れた場合には NP として扱う。それ以外の場合には型 NP&{〈時〉}を ←VP として扱うことで、(4), (5), (6) の yesterday に型 NP&{〈時〉} だけを与えて解析することができる。このように、前後の型からどちらとして働いているかを決定できるような2つの型は素性値に用いてまとめることができる。素性値を用いることで、基本型を増やさずに単語に与えられる型の数を抑えられることも語彙作用文法の特徴である。さらに、素性値は還元によって伝搬されるため、特定の語順などによって生じる統語的な規則も扱いやすいという利点がある。例えば、two などの数を表す基数名詞は →NP として後ろの名詞を修飾する働きがある。そして修飾された名詞は量を持った名詞、量名詞として普通の名詞とは異なる扱いを受けることがある。このような規則は還元の際に、基数名詞が基数を表す素性値〈CRD〉を後ろの名詞に渡すことによって実現することができる。

5 むすび

本論文では構文解析の際に、修飾子による解析木爆発を防ぐために、素性値によって修飾を制限し、文の核構造のみを生成する語彙作用文法を提案した。

また、左再帰自由と呼ばれる性質を有する英語の添え字付加言語が決定性言語であることを示した。

今後は、型の組み合わせによる解析木の増加を抑えるための方法や、修飾子の修飾先を決定する情報を収集して、核構造をよりもっともらしい構造にする方法の研究を行う必要があると考えている。

参考文献

[1] 川原田 郁夫, 笠井 琢美, “転送スタック付きプッシュダウンオートマトンと Linear Indexed Grammar について”, 電子情報通信学会論文誌, vol.J84-D-1,no.12,1583-1590, 2001.
[2] 鶴岡 友, “統語素性を利用した構文解析”, 電子情報通信学会技術研究報告, vol.105, no.680, COMP2005-65, 15-21, 2006.
[3] 松原俊一, 笠井琢美, “拡張範疇文法—機械翻訳のための新しい文法モデル—”, 電子情報通信学会論文誌, vol.J92-D, no.9, 1508-1517,2009.
[4] アイバンA.サグ,トーマス・ワソー（著）, 統語論入門 形式的アプローチ, 郡司隆男・原田康也（訳）, 岩波書店, 2001
[5] 安井 稔, 英文法総覧-改訂版-, 開拓社, 1983