

自然言語処理を考慮したインドネシア語文型パターンの作成とその応用の一考察†

Hartono†† 田中穂積††

計算機上で、自然言語処理システムを構築する場合に、文を効率よく、的確に処理する必要がある。この時、文型パターンが重要な役割を果たす。例えば、英語については Hornby の文型パターンがよく知られている。英語の処理では、文型パターンは文を解析する際に必要となるだけでなく、正しく文を生成するためにも重要な役割を果たしている。英語の文型に類似するインドネシア語についても同じことが言える。しかし、インドネシア語については、このような文型パターンはまだ考えられていない。これからの計算機上でのインドネシア語処理を考えれば、インドネシア語の文型パターンを整理することは早急な課題と言えよう。本稿では、計算機処理に用いることを前提としたインドネシア語の文型パターンを提案する。この文型パターンは、動詞を伴う文型と動詞を含まない文型からなり、文型パターンの中核となる動詞型については、動詞を修飾する前置詞パターンをも考慮して分類する。本稿で提案した文型パターンの応用としてインドネシア語の統語解析、意味解析、訳語選択および生成について考察する。

1. はじめに

1.1 本研究の背景と目的

自然言語の文（以下「文」と書く）は決められた規則に従って単語を列挙したものである。こうした規則は文法と呼ばれており、言語ごとに異なる。われわれは、文法に従って文を生成したり、文法を用いて文を理解したりする。計算機上で、自然言語処理システムを構築する場合に、機械翻訳システムであろうと、質問応答システムであろうと、文法と様々な知識を用いて文の意味を理解し、理解した結果を抽出することが大切である。しかし、文を効率よく、的確に計算機上で処理するためには、文型パターンが重要な役割を果たす。例えば、英語については Hornby の文型パターン¹⁾がよく知られている。Hornby の文型パターンを基にした自然言語処理システム上の英語文法¹²⁾もある。このような文型パターンは、文を解析する時に必要となるだけでなく、正しく文を生成するためにも役に立つ。

例えば、英語の「tell」と「explain」について考えてみよう。「なぜかを私に教えてくれ」という日本語の文を英語に訳すと、「tell me why.」になる。ところが、「なぜかを私に説明しなさい」という文を「explain me why.」と訳してはいけない。「explain me why.」は、文法的に正しい文ではないのである。

「tell」と「explain」の構文パターンを知っていれば、このような間違った文の生成を避けることができると Hornby が指摘する。

英語の処理では、文型パターンが重要な役割を果たす。英語の文型に類似するインドネシア語についても同じことが言える。しかし、こうしたインドネシア語の文型パターンはまだ考えられていないのである。これからの計算機上でのインドネシア語処理を考えれば、インドネシア語の文型パターンを整理することは早急な課題と言えよう。

1.2 インドネシア語の特徴

インドネシア語は接辞を有力な文法要素としていることから言語の形態的分類上からは、こう着語に属していると言われている。原形の語（以下「素語」と書く）が接辞の付加によってその素語の品詞および意味を変えることがある¹⁰⁾。例えば、接頭辞 di が素語 makan（食べる）に付加すると受け身の dimakan（食べられる）となり、素語 makan と接尾辞 an から名詞の makanan（食べ物）が構成される。

- makan（食べる）
- dimakan [接頭辞 di+素語 makan]（食べられる）
- makanan [素語 makan+接尾辞 an]（食べ物）

また、インドネシア語は時制を持たない。発話された文によって表されるイベントがいつ起きたかを伝える手段としては、文の中に時間的副詞を入れるかまたはアスペクトマーカを使う。例えば例1、2では、イベントの時間は過去にも未来にも取れるが、saya datang dari Indonesia tahun lalu（私は昨年インド

† An Indonesian Sentence Pattern and Its Application for Natural Language Processing by HARTONO and HOZUMI TANAKA (Department of Computer Science, Faculty of Engineering, Tokyo Institute of Technology).

†† 東京工業大学工学部情報工学科

ネシアから来た)では、時間的副詞 *tahun lalu* (昨年) でイベントが過去で起きたことを表す。Mereka telah mengirimkan saya buku (彼らは(すでに)本を送ってくれた)。アスペクトマーカ *telah* (「イベントがすでに終わったことを表す」) で文が過去であることを表現する。もし文にこれらの時間的副詞・アスペクトマーカなどが陽に表現されていなければ、その場合は文脈で判断する。

1. Saya datang dari Indonesia.

(i) (come) (from) (Indonesia)

[訳] (1) わたしはインドネシアから来た。

2. Mereka mengirimkan saya buku.

(they) (send) (i) (book)

[訳] (2) 彼らは(私に)本を送ってくれた。

例1と2からもわかるようにインドネシア語は英語の文型に類似するところが多い。しかし、動詞型(動詞の種類によって構成可能な文のパターン)によって文型の説明がつく英語に対して、文中に必ずしも動詞が存在するとは限らないインドネシア語は、動詞型だけでは文型を説明することができない。また、時間的副詞と場所的副詞と様態的副詞が文法的に強く束縛されていない点もインドネシア語の特徴の一つである。

1.3 本論文の構成

前節で述べた内容に関し、本論文では、次のような構成で議論を進める。

第2章では、インドネシア語の動詞について説明し、動詞型を提案する。第3章では、動詞型と動詞が存在しない文を含めてインドネシア語の文型を記述する。第4章では、名詞と動詞を修飾する前置詞パターンを記述する。第5章では、本稿で提案した文型パターンの応用としてインドネシア語の統語解析、意味解析、訳語選択および生成について考察する。第6章では、本論文の結論について述べる。

なお、本稿で使用したインドネシア語の例文は教科書、参考書、雑誌から取ったもの³⁾⁻⁹⁾である。

2. 動 詞

2.1 動詞の種類

接辞を有力な文法要素としているインドネシア語には三十種類以上の接辞がある。インドネシア語の接辞には接頭辞、接尾辞に加えて複合辞(接頭辞と接尾辞の組み合わせでできた接辞)がある。このうち動詞を構成する接辞には、接頭辞として *meN-*, *ber-*, *di*, *ter-*, *memper-*, *diper-* があり、接尾辞として *-kan*, *-i* があ

る。また複合辞には *meN-kan*, *meN-i*, *di-kan*, *di-i*, *ter-kan*, *ter-i*, *ke-an*, *peN-an*, *ber-an* がある。

接辞を考慮することにより動詞は二種類に分けられる。接頭辞 *meN-* が付加された動詞とそれ以外の動詞である。接頭辞 *meN-* を付加した動詞には、

- (1) 受動態に変換不可能かつ目的語を取らない自動詞(例3)。
- (2) 目的語を取り、受動態に変換可能な他動詞(例4)。
- (3) 自動詞または他動詞として利用可能な自他動詞(例5)。

がある。

3. Dia *menangis*. (彼は泣く)

4. Dia *membeli sayur*. (彼は野菜を買う)

5. Dia *membangun* (rumah). (彼は(家)を建てる)

接頭辞 *meN-* が付加されていない動詞はすべて受動態に変換不可能な自動詞である。素語のみの動詞と接辞(接辞 *meN-* を除く)が付加した動詞も自動詞と考え、補語を取るもの(例6)、取らないもの(例7, 8)、取っても取らなくてもよいもの(例9)に分類することができる。

6. Rumah itu *kemasukan pencuri*. (あの家は泥棒に入られた)

7. Dia *pergi*. (彼は行く)

8. Dia *kedingin*. (彼は寒さにやられた)

9. Dia *makan* (pisang). (彼は(バナナを)食べる)

2.2 動 詞 型

前節の分類のように受動態に変換可能と変換不可能の二種類の動詞を考慮すれば、例えば、例4と6をそれぞれ

- 主語+他動詞 VT+名詞, と
- 主語+自動詞 VI+名詞

のように動詞型を分けることができる。しかし、そのパターンを見れば、どちらも「主語+動詞+名詞」を成している。文法を説明する際に、上のような動詞の後に続く名詞の役割の違いを区別する必要があるが、文型だけを考えれば両方とも同じパターンである。本稿で扱う動詞型は上のような動詞の文法的特性(自動詞、他動詞など)を別々に考慮せずに分類する。

また、動詞の修飾可能な前置詞句のパターンを利用することによって動詞の意味的特性がわかる。これは、インドネシア語処理における解析、生成の際に役に立つだけでなく、訳語選択にも使えるので、本稿で

は、これらの前置詞句も考慮して動詞型を分類する。これによって得られた動詞型は以下で示す。なお、{ }は任意的なものである。

- 主語+KK*1+名詞(句)+{副詞句}
- 主語+KK 2 +名詞(句)+動詞(句)+{副詞句}
- 主語+KK 3 +名詞(句)+形容詞(句)+{副詞句}
- 主語+KK 4 +名詞(句)+名詞(句)+{副詞句}
- 主語+KK 5 +名詞(句)+節+{副詞句}
- 主語+KK 6 +疑問節+{副詞句}
- 主語+KK 7 +*bahwa* 節+{副詞句}
- 主語+KK 8 +{副詞句}
- 主語+KK 9 +形容詞(句)+{副詞句}
- 主語+KK 10+動詞(句)+{副詞句}

例として、例えば、動詞 *pergi* (行く) は、動詞型 **KK 8** と **KK 10** を持つほかに、以下のような前置詞句が修飾するものと規定する。なお、これらの前置詞句については4章で詳しく説明する。記号については、例えば *p 1 b* は4章の前置詞句パターン 1. (b) に対応する。

- p 1 b*: *dari* +名詞(句)→[場所]
- p 1 c*: *ke* +名詞(句)→[場所]
- p 1 d*: *kepada* +名詞(句)→[人間]
- p 1 f*: *sampai* +名詞(句)→[場所]
- p 2 b*: *dari* +名詞(句)→[時間]
- p 2 e*: *sampai* +名詞(句)→[時間]
- p 3 a*: *dengan* +名詞(句)→[乗物]
- p 4 a*: *dengan* +名詞(句)→[生物]
- p 4 b*: *tanpa* +名詞(句)→[具体物]
- p 4 c*: *sama* +名詞(句)→[生物]
- p 4 d*: *bersama*+名詞(句)→[生物]
- p 5 a*: *sebagai* +名詞(句)→[人間]
- p 8 a*: *dengan* +形容詞
- p 8 b*: *dengan* +動詞(句)
- p 8 d*: *tanpa* +動詞(句)
- p 9 b*: *bagi* +名詞(句)→[具体概念]
- p 9 d*: *buat* +名詞(句)→[具体概念]
- p 9 a*: *demi* +名詞(句)→[任意]
- p 9 e*: *untuk* +名詞(句)→[任意]
- p 9 f*: *untuk* +動詞(句)

これによって、動詞 *pergi* で表現できる文は例えば、
10. *Dia pergi dengan isterinya ke Jakarta pada hari minggu untuk mengunjungi familinya.*

* **KK** は *Kata Kerja* (動詞) の略である。

** **BK** は *Bentuk Kalimat* (文型) の略である。

- p 4 a*: *dengan isterinya* (彼の奥さんと一緒に)
- p 1 c*: *ke Jakarta* (ジャカルタへ)
- p 2 d*: *pada hari minggu* (日曜日に)
- p 9 f*: *untuk mengunjungi familinya*
(家族を訪れるために)

【訳】(10) 彼は日曜日に奥さんと一緒に家族を訪問するためにジャカルタに行った。

また、インドネシア語の副詞について見てみると、出現する場所は文法的に強く束縛されていない(例 11~16)。

11. *Dia pergi ke Jakarta kemarin.*
(彼) (行く) (へ) (ジャカルタ) (昨日)
12. *Kemarin dia pergi ke Jakarta.*
13. *Dia kemarin pergi ke Jakarta.*
14. *Ke Jakarta dia pergi kemarin.*
15. *Kemarin ke Jakarta dia pergi.*
16. *Ke Jakarta kemarin dia pergi.*

【訳】(11)~(16) 彼は昨日学校に行った。

このように、副詞句の出現場所が任意的ではあるが、一般には動詞型の中で示した通り(文の最後)である。また、複数個あった場合には、並べる順序は原則的に任意である。

3. 文 型

英語の文には動詞が必ず存在する。したがって動詞型だけで英語の文型を説明できる。しかし、インドネシア語の文には必ずしも動詞があるとは限らない。

17. *Dia guru.*
(He/She) (teacher)
18. *Leni cantik.*
(Leni) (pretty)
19. *Pak Suzuki ke Indonesia.*
(Suzuki) (to) (Indonesia)

【訳】(17) 彼は先生だ。(18) レニさんはきれいだ。(19) 鈴木さんはインドネシアに(行った)。

例 17, 18 の文型はそれぞれ「名詞+名詞」と「名詞+形容詞」で、英語では *be* 動詞を用いて文を構成する。また、例 19 の場合は動詞が省略されている。この文を英語か日本語に訳したければ動詞が必要である。

以下では、第2章で説明した動詞型と上で述べた動詞のないインドネシア語の文型¹¹⁾とを合わせて記述する。

● **BK**1** は、動詞 **KK 1** の後に名詞句が続くパターンである。

なお、例 24 はこの型には入れない。(自動詞+名詞)ではあるが、この文の naik haji (巡礼に行く) は動詞+名詞と考えるよりは自動詞の熟語と扱うのが一般的である。したがって、例 24 は BK 8 に入れる。

BK 1

主語+KK 1	名詞(句)
20 Dia sering membaca	majalah.
21 Dia membeli	buku kemarin.
22 Joni naik	sepeda.
23 Ini merupakan	langkah penting.
24 Dia naik	haji.

【訳】(20) 彼はよく雑誌を読む。(21) 彼は昨日日本を買った。(22) ジョニさんは自転車にのる。(23) これは大事なステップのようだ。

● BK 2 は KK 2 の目的語となる名詞(句)の後に動詞(句)が続く文型である。

BK 2

主語+KK 2	名詞(句)	動詞(句)
25 Dia membawa	adiknya	pergi ke sekolah.
26 Dia membiarkan	adiknya	menangis.
27 Dia mengantarkan	temannya	pulang.

【訳】(25) 彼は弟さんを学校に連れて行った。(26) 彼は弟さんが泣いているのをそのままにしておく。(27) 彼は友達を家まで送った。

● BK 3 は KK 3 の目的語となる名詞(句)の後に形容詞が続く文型である。

BK 3

主語+KK 3	名詞(句)	形容詞
28 Ia mencat	rumahnya	merah.
29 Kami mengharapkan	abangnya	baik-baik.
30 Adik membuat	mereka	senang.

【訳】(28) 彼は自分の家を赤く塗った。(29) 私たちは兄さんが元気であるように望んでいる。(30) 弟さんは彼らを楽しんだ。

● BK 4 は動詞 KK 4 の後に名詞(句)が二つ続く文型である。

BK 4

主語+KK 4	名詞(句)	名詞(句)
31 Pak Tanaka menamakan	anaknya	Tom.
32 Mereka memanggil	saya	Har.
33 Dia mengiriskan	saya	buku.
34 Kakek itu memberikan	cucunya	boneka itu.

【訳】(31) 田中さんは子供にトムと名付ける。(32)

彼らはわたしをハルと呼ぶ。(33) 彼はわたしに本を送ってくれた。(34) おじいさんは孫に人形をあげた。

● BK 5 は KK 5 の目的語となる名詞(句)の後に節が続く。

BK 5

主語+KK 5	名詞(句)	節
35 Dia meyakinkan	bossnya	bahwa itu benar.
36 Dia menasehati	temannya	supaya tidak bolos.

【訳】(35) 彼はそれが正しいと上司に確信させた。(36) 彼はさほらないよう友達にアドバイスした。

● BK 6 は疑問節が動詞 KK 6 の後に続く文型である。

BK 6

主語+KK 6	疑問節
37 Mereka memperdebatkan	kemana seharusnya dia pergi.
38 Kita harus menentukan	kapan ia boleh keluar.
39 Mereka bertanya	kapan dia akan tiba.

【訳】(37) 彼らは彼がどこに行くべきかを討論した。(38) われわれはいつでられるかを決めなければならない。(39) 彼らはいつ彼が到着するかと聞いた。

● BK 7 は bahwa 節を動詞型 KK 7 の目的語として取る文型である。

BK 7

主語+KK 7	bahwa 節
40 Mereka menyangka	bahwa saya orang jawa.
41 Dia menganggap	bahwa soal itu beres.

【訳】(40) 彼らはわたしがジャワ人だと思っている。(41) 彼はあの問題が解決したものだ判断した。

● BK 8 は述部に補語を伴わずに動詞型 KK 8 のみで構成される文型である。この型には、tidur (寝る) のような自動詞と、場合によっては目的語を取れる makan (食べる) のような自他動詞が含まれている。

BK 8

主語+KK 8
42 Dia menangis.
43 Amat sedang makan.
44 Anak-anak berdiri.

[訳](42) 彼は泣く。(43) アマッさんは食べている。(44) こどもたちは起立する。

● BK 9 は動詞 KK 9 後に形容詞が続く文型である。この形容詞は結果を指示する主補語として用いられる。

BK 9

主語+KK 9	形容詞
45 Dia kawin	muda.
46 Bunga ini berbau	harum.
47 Rumahnya bercat	putih.

[訳](45) 彼は若くして結婚した。(46) この花はいいにおいがする。(47) 彼の家は白いペンキで塗られている。

● BK 10 は動詞 KK 10 の後に動詞(句)が続く文型である。

BK 10

主語+KK 10	動詞(句)
48 Ibu Suzuki datang	memanggil anaknya.
49 Dia tersenyum	melihat saya.
50 Amat disuruh	pulang ke kampung.
51 Dia terkejut	mendengar berita itu.

[訳](48) 鈴木さんは子供を呼びにきた。(49) 彼はわたしを見て微笑んだ。(50) アマッさんは田舎に帰るように言われた。(51) 彼はあのニュースを聞いて驚いた。

● BK 11 は述部として名詞(句)を取るが、主語と述語の間に adalah/ialah が入ることもある。

BK 11

主語	(adalah/ialah)+名詞(句)
52 Ayah temannya	dosen Universitas Indonesia.
53 Orang itu	anak Gubernur.
54 Anak-anak itu	penduduk kota Medan.

[訳](52) 友達のお父さんはインドネシア大学の教授だ。(53) あの人は知事の息子だ。(54) あの子どもたちはメダンの市民だ。

● BK 12 は主語の後に前置詞句が続く文型であるが、これは KK 8 の動詞を省いたものだと考えてもよい。

BK 12

主語	前置詞	名詞(句)
55 Saya	dari	Sumatra Utara.
56 Cincin ini	untuk	kamu.
57 Anak-anak itu	dengan	abangnya.

[訳](55) わたしは北スマトラから(来た)。(56) この指輪は君に(買ってあげた)。(57) 子どもたちは兄さんと(一緒にいる)。

● BK 13 は述部として形容詞(句)を取る文型である。

BK 13

主語	形容詞(句)
58 Pakaiannya	sangat sederhana.
59 Rumah Pak Tanaka	megah.
60 Mukanya	pucat.

[訳](58) 彼女の服はたいへん簡素だ。(59) 田中さんの家は立派だ。(60) 彼の顔色は青ざめた。

● BK 14 は述部として形容詞と名詞を持つ。

BK 14

主語	形容詞	名詞
61 Dia	sakit	kepala.
62 Wanita itu	panjang	lidah.

[訳](61) 彼は頭が痛い。(62) あの女はおしゃべりだ。

ただし、例 62 はこの文型に入れない。panjang lidah (おしゃべり) はもはや分けることのできない語いであり、名詞の熟語として処理する。したがって例 62 のような文は BK 11 型に入る。

● BK 15 は述部として形容詞に続いて動詞(句)が並ぶ文型である。

BK 15

主語	形容詞	動詞(句)
63 Dia	berani	makan ikan mentah.
64 Ibunya	suka	berjalan ke luar negeri.

[訳](63) 彼は生魚を食べる勇気がある。(64) お母さんは海外旅行が好きだ。

● 述部として形容詞の後に bahwa 節が続く文型は BK 16 である。

BK 16

主語	形容詞	bahwa 節
65 Dia	gembira	bahwa anaknya lulus ujian.
66 Adi	kecewa	bahwa temannya membohonginya.

[訳](65) 彼は子供が試験に合格したことを喜んでいる。(66) アディさんは友達にだまされたことに悔しがつている。

● BK 17 は単語 ADA (存在を表す) を先頭に

名詞(句)が続く文型である。

BK 17

Ada	名詞(句)
67 Ada	kecelakaan mobil kemarin.
68 Ada	gempa bumi terjadi di Tokyo.

[訳](67) 昨日事故があった。(68) 東京で地震が起きた。

4. 前置詞句

本章では、格の意味的特性を考慮し、前置詞句パターンを提案する。以下では、その一部分を示す。なお、パターンの記号として、例えば、1aの場所・方向を示す「di+名詞句」は **pla** で表すとする。

1. 場所・方向の指示

- (a) di+名詞(句) (～[場所] で)
- (b) dari+名詞(句) (～[場所] から)
- (c) ke+名詞(句) (～[場所] へ)
- (d) kepada+名詞(句) (～[場所] に)
- (e) pada+名詞(句) (～[場所] で)
- (f) sampai+名詞(句) (～[場所] まで)

2. 時間の指示

- (a) di+名詞(句) (～[時間] に)
- (b) dari+名詞(句) (～[時間] から)
- (c) hingga+名詞(句) (～[時間] まで)
- (d) pada+名詞(句) (～[時間] に)
- (e) sampai+名詞(句) (～[時間] まで)
- (f) sejak+名詞(句) (～[時間] から)
- (g) semenjak+名詞(句) (～[時間] から)
- (h) waktu+名詞(句) (～[時間] に)
- (i) takkala+名詞(句) (～[時間] に)
- (j) sesudah+名詞(句) (～[時間] の後に)
- (k) setelah+名詞(句) (～[時間] の後に)
- (l) sebelum+名詞(句) (～[時間] の前に)
- (m) selama+名詞(句) (～[時間] の間に)

3. 使用道具の指示

- (a) dengan+名詞(句) (～[道具] で)
- (b) tanpa+名詞(句) (～[道具] なしで)

4. 同伴の指示

- (a) dengan+名詞(句) (～[生物] と一緒に)
- (b) tanpa+名詞(句) (～[生物] 抜きで)
- (c) sama+名詞(句) (～[生物] と一緒に)
- (d) bersama+名詞(句) (～[生物] と一緒に)

5. 役割の指示

- (a) sebagai+名詞(句) (～として)

6. 類似の指示

- (a) seperti+名詞(句) (～のように)
- (b) laksana+名詞(句) (～のように)
- (c) sebagai+名詞(句) (～のように)

7. 内容の指示

- (a) akan+名詞(句) (～[物事] について)
- (b) tentang+名詞(句) (～[物事] について)
- (c) mengenai+名詞(句) (～[物事] について)
- (d) perihal+名詞(句) (～[物事] について)

8. 様態の指示

- (a) dengan+形容詞 (～的に)
- (b) dengan+動詞(句) (～することで)
- (c) secara+形容詞 (～的に)
- (d) tanpa+動詞(句) (～しないで)

9. 目的の指示

- (a) demi+名詞(句) (～のために)
- (b) bagi+名詞(句) (～のために)
- (c) guna+名詞(句) (～のために)
- (d) buat+名詞(句) (～のために)
- (e) untuk+名詞(句) (～のために)
- (f) untuk+動詞(句) (～するために)

69. Dia pergi **kemarin**.

70. Dia pergi *pada hari minggu*.

71. Dia pergi **cepat-cepat**.

72. Dia pergi *dengan sedih*.

[訳](69) 彼は昨日行った。(70) 彼は日曜日に行った。(71) 彼は速く行った。(72) 彼は悲しく去って行った。

kemarin (昨日) (例 69) と **cepat-cepat** (速く) (例 71) のような一部の時間的副詞と様態的副詞を除いてインドネシア語の副詞はすべて前置詞句によって表される (例 70, 72)。2.2 節で述べたようにインドネシア語の副詞の出場所自由である。しかし、これによって曖昧さを生じる場合もある。

73. Dia pergi melihat bangunan *di Jakarta*.

74. Dia pergi melihat bangunan *ke Jakarta*.
(彼) (行く) (見る) (建物) (ジャカルタ)

[訳](73) 彼はジャカルタの建物を見に行った。(74) 彼はジャカルタに建物を見に行った。

例 73～74 の文のパターンはどちらも「名詞1+動詞1+動詞2+名詞2+前置詞句」である。この前置詞句は動詞1を修飾するか、名詞2を修飾するか両方考えなければならないが、2.2 節で記述した動詞 *pergi* の修飾可能な前置詞句を利用すれば、前置詞句

p1a が動詞1を修飾することがないから動詞 *pergi* と前置詞句 *di Jakarta* を深く考慮しなくてすむ (例 73).

また、前置詞句 p1c は方向を示す名詞を修飾することができる。方向を指示しない *bangunan* (建物) を修飾せずに、動詞 *pergi* だけを修飾することになる (例 74).

このように名詞、動詞に修飾可能な前置詞句で規定することによって無駄な意味の解釈を避けることができる。

5. インドネシア語処理における文型パターンの応用

5.1 解析と生成

統語解析の段階において動詞型の分類を適用することによって不必要な処理を避けることができる。

- a. sdec→np, v, np.
- b. sdec→np, v, np, adjp.
- c. sdec→np, v, np, np.

動詞の分類をせずに、そのまま処理すると、例えば、規則 c のようなパターンにしか適用しない動詞でも、規則 a, b が適用される。動詞型を利用することによって規則 a, b, c の代わりに以下に示す a', b', c' を用意することにより、このような無駄な処理を排除することができる。

- a'. sdec→np, KK 1, np.
- b'. sdec→np, KK 3, np, adjp.
- c'. sdec→np, KK 4, np, np.

また、冗長な (誤った) 構文木を排除することもできる。第4章の例 73 を見てみよう。

第4章で述べたように「名詞1 + 動詞1 + 動詞2 + 名詞2 + 前置詞句」の場合では、前置詞句が動詞1と名詞2のいずれかを修飾する。これによって構文木が二つ作られる。しかし、例 73 の前置詞句の場合には *bangunan* (建物) のみを修飾する。このような処理は動詞 *pergi* の修飾可能な前置詞句パターンを適用することによってその前置詞句が動詞 *pergi* を修飾しないことがわかるので解析の段階でこれを排除することができる。

生成においても、正しい文を生成するためには、文型パターンが役に立つことを 1.1 節で説明した。

5.2 意味解析

意味を適切に解析するためにも、本稿で提案する動詞型が役に立つ。

75. Dia mengirimkan buku kepada saya.

76. Dia mengirimkan saya buku.

(彼) (送る) (私) (本)

[訳](75)(76) 彼は私に本を送った。

例 75 と 76 の動詞 *mengirimkan* の動詞型はそれぞれ **KK 1** と **KK 4** である。例 75 の動詞 *mengirimkan* (送る) の後に続く名詞句 *buku* (本) が OBJECT (対象格) であるに対して例 76 では、同じ動詞 *mengirimkan* の後に続く名詞句 *saya* (私) は RECIPIENT (権利終点) である。動詞型 **KK 1** と **KK 4** を利用することによって、名詞句の存在する位置を知り、上の例のような同じ動詞の後に続く名詞句のそれぞれの意味を正確に捉えることができる。

5.3 訳語選択

また、機械翻訳における訳語選択にも文型パターンが役に立つ。例えば、

77. 彼は学校に本を持って行く。

78. Dia *membawa* buku ke sekolah.

79. ★Dia *mengambil* buku ke sekolah.

80. 彼は家から本を持って行く。

81. Dia *mengambil* buku dari rumah.

82. ★Dia *membawa* buku dari rumah.

「持って行く」がインドネシア語では、「*membawa*」または「*mengambil*」に訳すことができる。同じ「本を持って行く」であるから例 77 と例 80 はそれぞれ例 78, 79 と例 81, 82 に訳すことができるであろう。しかし、正しい訳文はそれぞれ例 78 と例 81 だけである。これらの例のように、動詞の対象格の意味的特性だけをを用いることでは正しく訳し分けることができない。しかし、文型パターンにより、動詞 *mengambil* は、前置詞句 **p1c** を取らないパターンであることを知り、また動詞 *membawa* は、前置詞句 **p1c** を取り、一般に **p1c** と一緒に出現しなければ、単独に **p1b** を修飾しないことがわかるので、文を正しく訳し分けることが可能となる。

6. おわりに

本稿では、コンピュータにおけるインドネシア語処理を目的としてインドネシア語の文型パターンを提案した。この文型パターンの中核となる動詞型については、動詞を修飾可能な前置詞句パターンをも考慮して分類したものである。

また、インドネシア語の統語解析、意味解析、訳語選択および生成における本稿で提案した文型パターン

の応用についても考察した。

付加疑問文、疑問文、受身文などの変形による文型は、本稿で提案した文型には入っていない。今後の課題として、これらの文型をも含めて、動詞周辺に付加するアスペクトマーカなどについても分類の基準として考慮する必要がある。また、より多くの例文を参照したり、この文型パターンを用いた自然言語処理システムを作成し、文型パターンの改善を行う予定である。

参 考 文 献

- 1) Hornby, A. S.: *Oxford Advanced Learner's Dictionary of Current English Third Edition*, The English Language Book Society and Oxford University Press (1974).
- 2) Bambang Kaswanti Purwo: *Deiksis Dalam Bahasa Indonesia*, PN. Balai Pustaka (1984).
- 3) Surana dkk, F. X.: *Himpunan Materi TATA BAHASA I, II, III*, Penerbit Tiga Serangkai (1983).
- 4) Tarigan, H. G.: *Pengajaran KOSAKATA*, Penerbit ANGKASA Bandung (1985).
- 5) Jos Daniel Parera: *Bidang Sintaksis*, Penerbit Nusa Indah (1982).
- 6) Ramlan, M.: *SINTAKSIS*, UP. Karyono Yogyakarta (1981).
- 7) Samsuri: *Tata Kalimat BAHASA INDONESIA*, Sastra Hudaya (1982).
- 8) Soenjono Dardjowidjojo: *Indonesian Syntax*, Ph. D. Thesis, Georgetown University (1967).
- 9) Wojowasito, S.: *Pengantar Sintaksis Indonesia*, Penerbit Sintha Dharma, Bandung

(1976).

- 10) Hartono, 田中: Lang LAB 上でのインドネシア語文法試作, 第 33 回情報処理学会全国大会論文集, pp. 1433-1434 (1986).
- 11) Hartono, 田中: 自然言語処理を考慮したインドネシア語文型パターン, 情報処理学会自然言語処理研究会資料 68-5 (1988).
- 12) 岩山, 徳永, Quek, 田中: 自然言語処理のための英語文法, 第 37 回情報処理学会全国大会論文集, pp. 1100-1101 (1987).

(昭和 63 年 12 月 22 日受付)
(平成 元年 4 月 11 日採録)



Hartono (正会員)

1962 年生. 1986 年東京工業大学工学部情報工学科卒業. 1988 年同大学院修士課程修了. 現在, 同大学院博士課程在学中. 機械翻訳に関する研究に従事. 日本ソフトウェア科学会, 認知科学会各会員.



田中 穂積 (正会員)

昭和 39 年東京工業大学理工学部制御工学科卒業. 昭和 41 年同大学修士課程修了. 同年電気試験所(現, 電子技術総合研究所)入所. 昭和 58 年東京工業大学工学部情報工学科助教授. 昭和 61 年同大学教授となり現在に至る. 工学博士. 人工知能, 自然言語処理の研究に従事. 電子情報通信学会, 認知科学会, 日本ソフトウェア科学会, 人工知能学会, 計量国語学会各会員.