

述語取り立て形式の整理と文生成器への実装

佐藤理史

名古屋大学大学院工学研究科情報・通信工学専攻

ssato@nuee.nagoya-u.ac.jp

1 はじめに

文の内部表現から文字列としての文を生成するソフトウェアは、一般に、文生成器 (surface realizer) と呼ばれる。我々は、日本語文生成器である HaoriBricks[1] を、日本語テキストを生成するためのドメイン固有言語 (domain specific language; DSL) と位置づけ、小説の自動生成の使用に耐えるように機能強化を進めている。

一万字近い短編小説の文章を生成するプログラムをできるだけ簡便に記述するためには、コードの再利用性を高める必要がある。実際、一つの小説において、まったく同じ文というものは、ほとんど存在しない。異なる文を生成するためには異なるコードが必要であるから、文の種類だけコードを書かなければならない。そのため、文を生成するコードをできるだけコンパクトに記述できることが重要となる。その一つの方法は、パラメータを利用したコード共有化 (つまり、サブルーチン化) であり、これにより、同じ文型の文をサブルーチン名とパラメータの記述のみで生成できるようになる。HaoriBricks には、この機能が最初から組み込まれている。

もう一つの方法は、いわゆる構文的操作や変形操作のサポートである。このような操作を実装すれば、基本となる文から、その否定文、疑問文、強調文など、文型の変更を伴うバリエーションを簡便に生成できるようになる。テンスの追加、末尾への「ない」の付与、終助詞の付与など、何らかの要素を文構造の「外」に追加する操作の実装は比較的容易であり、ブロックモデルの採用によって自然な形で取り込まれている。これに対して、文構造の「中」に手を伸ばして、その一部を書き換える操作の実装は、それほど単純ではない。

本稿では、このような変形操作の一つとして、取り立て助詞による述語の取り立てを取り上げる。「取り立て」は、「文のある要素をきわだたせ、同類の要素との関係を背景として、特別な意味を加える [3]」操作であり、日本語文法の主要な項目の一つである。取り

立ての中心的な役割を果たすのは、「も」「は」「なら」「だけ」「しか」「ばかり」「こそ」「さえ」「まで」「でも」「だって」「なんか」「なんて」「など」「くらい」などの取り立て助詞であり、それらが格成分 (補足語)、副詞的成分 (連用修飾語)、述語、節などを取り立てる。

取り立て助詞による述語の取り立て例を以下に示す。

- (1) a. 鈴木さんの言葉はさすががしかった
b. 鈴木さんの言葉はさすががしくさえあった
- (1a) から (1b) を作り出す操作を、「取り立て助詞『さえ』で述語『さすががしい』をとりたてる操作」とみなす。本稿の目標は、(1b) を生成するコードを、(1a) を生成するコードを再利用して記述できるようにすることである。

2 述語取り立ての形式的整理

実装に先立ち、述語の取り立てを形式的に整理する必要がある。ここでは、『基礎日本語文法』[2] および『現代日本語文法 5』[3] に記述されている述語の取り立てを、次の 3 種類に大別して整理する。

2.1 形式動詞の付加を必要とする形式

形式動詞「する」「ある」の付加を必要とする形式を以下に示す。

- (2) a. 書く (動詞型)
b. 書きもする
(基本連用形 + 取り立て助詞 + する)
- (3) a. 美しい (イ形容詞型)
b. 美しくもある
(基本連用形 + 取り立て助詞 + ある)
- (4) a. 確実だ (ナ形容詞型)
b. 確実でももある
(タ系連用テ形 + 取り立て助詞 + ある)
- (5) a. 本だ (判定詞型)
b. 本でももある
(タ系連用テ形 + 取り立て助詞 + ある)

述語の分類は、品詞ではなく活用型に基づく。たとえば、複合述語「書いてある」は動詞型で活用するので、「書いてありもする」というとりたて形式が可能である。ナ形容詞や判定詞のいわゆるデアル系列も「ある」が動詞型で活用するので、「確実にありもする」「本でありもする」という形式が可能である。

動詞型には「する」が付加されるのに対し、形容詞型と判定詞型（「だ」）には「ある」が付加される。複合述語「書きにくい」「書きそうだ」は形容詞型で活用するので、「書きにくくもある」「書きそうでもある」の形式が可能である。

2.2 形式動詞の付加を必要としない形式

形式動詞の付加を必要としない形式を以下に示す。

- (6) a. 書いている（タ系連用テ形 + テ助動詞）
b. 書いてもいる
- (7) a. 美しくあれ（基本連用形 + ある）
b. 美しくもあれ
- (8) a. 確実にある（基本連用形 + ある）
b. 確実でもある
- (9) a. 本である（基本連用形 + ある）
b. 本でもある

典型的には、動詞型の述語のタ系連用テ形と、それに接続する助動詞（本稿では、テ助動詞と呼ぶ）の間に取り立て助詞が挿入される形式である。さらに、ナ形容詞と判定詞のデアル系列においては、「で」と「ある」の間に取り立て助詞の挿入が可能である。さらに、イ形容詞に「ある」が付加された形は、「ある」の直前に取り立て助詞が挿入可能である。

2.3 否定が絡む形式

否定が絡む形式は、否定と取り立てを同時に考えるか、独立に（2ステップで）考えるかという、2つの立場がありうる。

- (10) a. 書く
b. 書かない（a. の否定）
c. 書きもしない
d. 書かなくもない
e. 書くしかない
- (11) a. 美しい
b. 美しくない（a. の否定）
c. 美しくもない
d. 美しくなくもない
e. 美しくしかない

- (12) a. 確実だ
b. 確実でない（a. の否定）
c. 確実でもない
d. 確実でなくもない
e. 確実でしかない
- (13) a. 本だ
b. 本でない（a. の否定）
c. 本でももない
d. 本でなくもない
e. 本でしかない

ここでは、第3の立場、すなわち、b. から c. への変形を「取り立て」、a. から c. への変形を「否定取り立て」として、どちらも考える立場をとる。上記の d. は、形式的には、否定形 (b.) の否定取り立てとみなせるので、暫定的にそのような扱いとする¹。

上記の e. は、特別なとりたて形式（「しかない」の付加）とみなす。これは、a. から e. の変換のみを考え、b. から e. への変換を考えない。なお、動詞型活用の場合、取り立て助詞の直前が基本形となる点が例外的である。

2.4 述語の取り立てとみなさい形式

『現代日本語文法5』[3]には、以下の例文がある。

- (14) 就職なんか/なんて/などしたくない

これは、名詞+「する」の間に取り立て助詞が挿入されるという点で例外的である。このような形式は、文献[3]では、「なんか」「なんて」「など」に対してのみ記述されている。

この文の元の形として、次の2種類が考えられる。

- (15) a. 就職したくない
b. 就職をしたくない

前者であれば、述語の取り立てとみなさざるを得ない。しかし、後者であれば、格成分の取り立てとみなすことができる。本稿では、暫定的に(14)の元の形式を(15b)とみなし、述語の取り立てとはみなさないこととする。

格成分（補足語）や副詞的成分（連用修飾語）の取り立ては、上記のような例を除けば、述語の取り立てと区別することは比較的たやすい。例文(16c)も、述語の取り立てとはみなさない。

- (16) a. 本だけを書く（格成分）

¹将来的には、二重否定の形式として整理する方が適切であろう。

- b. 足が痛くてゆっくりとしか歩けない
(副詞的成分)
- c. 持っていたのは本だけだ
(「本だけを持っていた」の「～は～だ」構文)

節の取り立ては、述語の取り立てとは区別する。さらに、述語を名詞化する取り立ても、節を作る操作の一部と考え、述語の取り立てとは区別する。

- (17) a. 困ったときだけ、神頼みする (節)
- b. お湯を注ぐだけでスープができる
(述語を名詞化)

なお、節の取り立ては、取り立て助詞を削除しても文として成立する。しかし、述語を名詞化する取り立ては、取り立て助詞を削除すると文として成立しない。

- (18) a. 困ったとき、神頼みする
- b. *お湯を注ぐでスープができる
- c. お湯を注 { ぐと/げば } スープができる

この他に、判断に迷う次のような形式がある。これらは、本稿では検討の対象外とする。

- (19) a. あとは報告書を書くだけだ (名詞化?)
- b. さっきから泣いてばかりだ (名詞化?)

3 HaoriBricks の取り立て操作

以上のような検討に基づき、以下に示す3種類の述語の取り立て操作(変形操作)を設定する。なお、HaoriBricksでは、形式動詞「する」「ある」は、一括して連用助動詞(直前に連用形またはテ形を要求する助動詞)として扱う²。

- (20) a. 述語
- b. 述語 + 取り立て助詞 + 連用助動詞する・ある
- c. 述語(動詞型) + しか + イ形容詞ない
- d. 述語(それ以外) + しか + 連用助動詞ある + 接尾辞ない
- (21) a. 述語 + テ・連用助動詞
- b. 述語 + 取り立て助詞 + テ・連用助動詞
- c. 述語 + しか + テ・連用助動詞 + 接尾辞ない
- (22) a. 述語
- b. 述語 + 取り立て助詞 + 連用助動詞する・ある + 接尾辞ない

²他に、形容詞の連用形に接続する「なる」(例: 美しくなる)、ナ形容詞のタ系連用形に接続する「いる」(例: 健康でいる)がある。

取り立てる前の形式(a.)の末尾の述語(ブロック体)に着目すると、取り立て助詞は、その述語の直後に挿入されるか、あるいは、直前に挿入される。前者(20)を「外(そと)取り立て」、後者(21)を「内(うち)取り立て」と名付ける。内取り立てが可能なのは、述語が複合述語³で、かつ、末尾の述語がテ助動詞か連用助動詞の場合に限られる。「否定取り立て」(22)は、外取り立ての一種で、取り立て助詞の挿入と「接尾辞ない」の付加を同時に行う操作である⁴。

取り立て助詞が「しか」の場合((20c),(20d),(21c))は、同時に「ない」を付加する。述語が動詞型の外取り立て(20c)は、連用助動詞の付加を必要としないという点で例外となる(そのゆえ、動詞は基本形となる)。

4 HaoriBricks における実装

HaoriBricksでは、あらかじめ用意されたブロック(関数)を組み合わせて、テキストを生成するコードを記述する。コード例を図1に示す。この図のtは、テスト用の関数で、第1引数に文字列として与えられたコードが、第2引数の文字列を生成することを表す。述語の取り立て操作に対応する3種類のブロックは、いずれも2引数の関数で、第1引数に取り立て助詞、第2引数に取り立ての対象となる構造(を生成するコード)をとる。

図1の26行目が、本稿の冒頭で示した例文(1b)を生成するコードである。このコードは、例文(1a)を生成するコード(24行目)に、「述語外取り立て」ブロックを追加しただけのコードとなっている。つまり、例文(1a)を生成するコード(x)がすでに存在していれば、「述語外取り立て(さえ, x)」とだけを書けば、例文(1b)を生成できる。

HaoriBricksは、ユーザーが記述したコードをHaori構造(構文構造)に一旦変換し、その後、表層文字列を生成する。取り立てブロックは、第2引数から作られるHaori構造において、最も後方にある述語を捕まえ、その述語に対して指定された取り立て操作を適用する。その模式図を図2に示す。外取り立ての場合は、図に示したような助動詞構造以下の構造を作成し、述語(「すがすがしい」)をその構造で置換する。この操作により、「～さえある」が挿入される。

「すがすがしい」を「すがすがしく」に活用する処理は、連用助動詞「ある」に対して定義されている接続型によって行われる。接続型とは、直前の形態素が

³「肌が美しくなる」なども複合述語とみなす。

⁴「連用助動詞ある」+「接尾辞ない」は、「ない」と表層化される。

```

1 t(' 述語外取り立て(も, 述語(動詞(:書く)))', '書きもする'),
2 t(' 述語外取り立て(も, ない(述語(動詞(:書く))))', '書きもしない'),
3 t(' 述語外取り立て(も, 述語(形容詞(:美しい)))', '美しくもある'),
4 t(' 述語外取り立て(も, ない(述語(形容詞(:美しい))))', '美しくない'), # 取り立てない
5 t(' 述語外取り立て(も, 述語(形容詞(:確実)))', '確実でもある'),
6 t(' 述語外取り立て(も, 述語(だ(:本)))', '本でもある'),
7 t(' 述語外取り立て(も, タ(述語(動詞(:書く))))', '書きもした'),
8 t(' 否定取り立て(も, 述語(動詞(:書く)))', '書きもしない'),
9 t(' 否定取り立て(も, 述語(形容詞(:美しい)))', '美しくもない'),
10 t(' 否定取り立て(も, 述語(形容詞(:確実)))', '確実でもない'),
11 t(' 否定取り立て(も, 述語(だ(:本)))', '本でもない'),
12 t(' 述語内取り立て(も, テいる(述語(動詞(:書く))))', '書いてもいる'),
13 t(' 述語外取り立て(も, テいる(述語(動詞(:書く))))', '書いていもする'),
14 t(' 述語内取り立て(も, テいる(テみる(述語(動詞(:書く))))', '書いてみてもある'),
15 t(' テいる(述語内取り立て(も, テみる(述語(動詞(:書く))))', '書いてもみている'),
16 t(' 述語内取り立て(も, テほしい(述語(動詞(:書く))))', '書いてもほしい'),
17 t(' 述語内取り立て(も, ない(述語(形容詞(:美しい))))', '美しくもない'),
18 t(' 述語内取り立て(も, 連用ある(述語(形容詞(:確実))))', '確実でもある'),
19 t(' 述語内取り立て(も, 連用ある(述語(だ(:本)))', '本でもある'),
20 t(' 述語外取り立て(しか, タ(述語(動詞(:読む))))', '読むしかなかった'),
21 t(' 述語外取り立て(しか, 述語(形容詞(:美しい)))', '美しくしかない'),
22 t(' 述語内取り立て(しか, タ(テいる(述語(動詞(:読む))))', '読んでしかいなかった'),
23 t(' 述語内取り立て(しか, 連用ある(述語(だ(:本)))', '本でしかない'),
24 t(' タ(提題(は(の(:鈴木さん, :言葉)), 述語(形容詞(:さすががしい))))',
25   '鈴木さんの言葉はさすががしくさえあった'),
26 t(' 述語外取り立て(さえ, タ(提題(は(の(:鈴木さん, :言葉)), 述語(形容詞(:さすががしい))))',
27   '鈴木さんの言葉はさすががしくさえあった'),

```

図 1: HaoriBrick コードの例

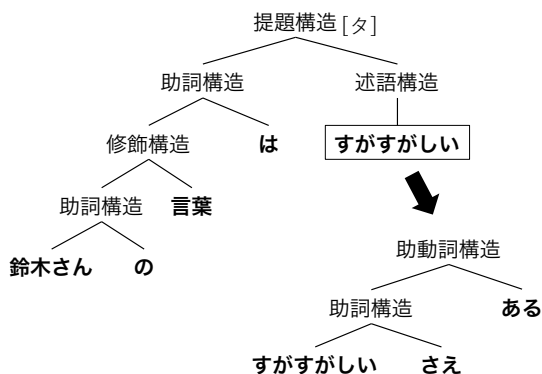


図 2: Haori 構造の書き換え (述語外取り立て)

どんな活用型の場合にどんな活用形を要求するかを類型化したものである。連用助動詞「ある」は、直前のイ形容詞に対して基本連用形を要求する。取り立て助詞は、この接続型による要求を透過するため、「ある」から「さすががしい」に基本連用形が要求され、表層形「さすががしく」が生成される。

テンスの情報は、Haori 構造の各節点を持つ属性リストに保持される。26 行目のコードでは、提題構造の外にテンスを付加する「タ」ブロックが結合されているため、テンス情報は提題構造ノードに保持される。

この情報は、構造の線形化の際に適切な葉ノード（この場合は「ある」）に伝達され、表層化の際に活用形としてタ形が選択され、「あった」が生成される。

HaoriBricks には、内取り立てが可能な場合は内取り立てを、不可能な場合は外取り立てを適用する「述語取り立て」ブロックも実装した。実際のコード記述では、ほとんどの場合、このブロックを用いればよく、内取り立てが可能で、かつ、外取り立てを行いたいときのみ(図 1 では 13 行目のみ)、「述語外取り立て」ブロックを用いることになる。

謝辞 本研究は JSPS 科学研究費挑戦的研究(萌芽)「ブロック玩具をモデルとする日本語文章生成ツールキットの設計と実装」(課題番号 17K20028)の助成を受けている。

参考文献

- [1] 佐藤 理史. HaoriBricks: ブロック玩具に学ぶ日本語文章生成ライブラリ. 言語処理学会第 23 回年次大会発表論文集, pages 20–23, 2017.
- [2] 益岡 隆志, 田窪 行則. 基礎日本語文法—改訂版—. くろしお出版, 1992.
- [3] 日本語記述文法研究会. 現代日本語文法 5 とりたて・主題. くろしお出版, 2009.