

言語教育力開発のための文法パーサー実装

A Grammar Implementation on a Parser for Language Teaching Developments

古賀弘毅、佐賀大学

KOGA, Hiroki, Saga University, hkoga@cc.saga-u.ac.jp

概要：本研修で、参加者は、自分が日本語初級前半を教えていると仮定し、教える文法項目「他動詞」、「対格」、「自由語順」を、Morgan, Jerry L.氏開発の構文解析器 Earley に実装する。Earley の配布資料を参考に、同構文解析器キットを自分のPCにコピーし、既習文法項目「自動詞」、「提題 I」、「主格」の実装例を掲載し、いくつか語列を入力して構文解析させ、使い方を知る。この後、新出の文法項目と一緒に議論しながら実装する。文法パーサー実装を通して、文法を系統立てて説明する仕方のヒントを得、学生の間違いや質問を予想する仕方を修得し、文法説明力の向上を図る。

キーワード：構文解析器への文法実装、提題詞「は」、自由語順、自・他動詞

1. 文法項目「自動詞」、「提題 I」、「主格」の実装例

図1は、教科書『みんなの日本語 I』の文法項目「自動詞」、「提題 I」、「主格」を、構文解析器 earley 用の文脈自由文法として実装した例である。

- | | | |
|---|--|-------------------------------------|
| 1. initial symbol: R | 2. $R \rightarrow \text{Topic } R/S$ | 3. $R \rightarrow \text{Subj } R/S$ |
| 4. $\text{Topic} \rightarrow N \text{ Top}$ | 5. $\text{Subj} \rightarrow N \text{ Nom}$ | 6. $\text{Top} \rightarrow wa$ |
| 7. $\text{Nom} \rightarrow ga$ | 8. $N \rightarrow kodomo$ | 9. $N \rightarrow gohan$ |
| 10. $N \rightarrow ojisan$ | 11. $N \rightarrow sakana$ | 12. $R/S \rightarrow nakimasu$ |
| 13. $R/S \rightarrow arukimasu$ | 14. $R/S \rightarrow nemasu$ | |

図 1: 日本語文法 1 : 自動詞、提題 I、主格形式

初期記号は「補語要求は満たされている関係 R」（「文」）である（規則 1）。自動詞は「主語を求める関係 R/S」とされる（規則 12 から 14）。「ga」は「主格形式 Nom」（規則 7）、「wa」は「提題詞 Top」（規則 6）とされ、それぞれ「名詞 N」をとって主語 Subj、提題 Topic とされる（規則 4、5）。自動詞 R/S は、主語 Subj か、提題 Topic か、どちらかを取って R となる（規則 2、3）。

文法は、もし語列が全体として R（文）であるなら、語列の統語構造を解析する。たとえば、日本語文法 1 は、語列「sakana wa nakimasu」を全体で R と分析し、図 2 のような統語構造を与える。もし語列が全体として R でないなら、解析なしである。たとえば、「nakimasu kodomo wa」は R ではない、つまり、非文である。

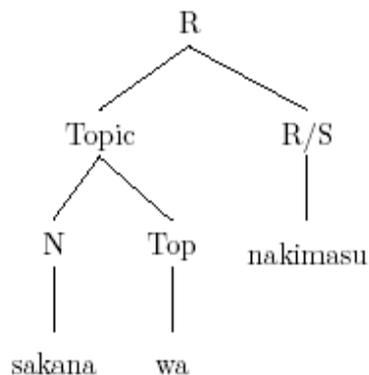


図 2: 日本語文法 1 による 'sakana wa nakimasu' の分析

2. 実装と文法説明力向上のヒント

2. 1 準備: 文法項目の射程の限定と次の射程の先読み

教える文法項目の射程には、どのような文や句の型が含まれるかを限定する。文法項目「他動詞」、「対格形式」、「自由語順」を教える場合、それにかかる時間は 2 時限で、その射程は以下 (1) 中のようなデータであろう。

- (1) a. こどもが さかなを たべます。
 b. さかなを こどもが たべます。
 c. こどもは さかなを たべます。
 d. *こどもが たべます さかなを。
 e. *さかなを たべます こどもが。

文 (1 a) は通常の SOV 文で、文 (1 b) は、主語と目的語との語順をかえた文である。文 (1 c) は SOV 文の主語が話題の文である。語列 (1 d) も、語列 (1 e) も、動詞が文末にない場合で、たとえ日常生活で聞くことがあっても、関係節の前や引用詞「と」の前では生起しないから、非文と考える。

同じ文法項目の次の射程に、どのような文や句の型があるかを先読みする。次の文法項目として「提題 2」と「自由語順 2」が考えられ、(2) 中のような話題を含み、語順がさまざまに変えられた文がその射程に入るであろう。

- (2) a. ?? さかなを こどもは たべます。
 b. さかなは こどもが たべます。
 c. ?? こどもが さかなは たべます。
 d. ? こどもは さかなは たべます。
 e. ? さかなは こどもは たべます

語列 (2 a) は、文 (1 c) の語順をかえたものであり、話題は、後方焦点で対比と解釈される。文 (2 b) は目的語が話題の文である。語列 (2 c) は、その話題と主語の語順

をかえた文である。文(2d)と(2e)とは、主語も目的語も話題となっている文で、お互い語順が異なるものである。

2. 2 説明力向上のヒント1：できる限り一般化する

当該の射程と先の射程を考慮し、説明をできる限り一般化しながら、当該の射程を説明する。これにより、系統だった文法説明が可能となり、また学生の質問や躓きも予想できる。日本語文法1で一般化した自動詞 R/S、 $R \rightarrow \text{Subj R/S}$ 、「主格形式 Nom」、「主語 Subj」に関する説明を参考にして、1) 他動詞は、主語と目的語を求める関係 R/SO であり(規則 14a から 14c)、2) 他動詞 R/SO は目的語 Obj を取って R/S となり(規則 3b)、3) 目的語 Obj は、対格形式 Acc「を」が名詞 N を取っているものとする(規則 7a、5a)。

3a. $R \rightarrow \text{Obj R/O}$	3b. $R/S \rightarrow \text{Obj R/SO}$
3c. $R/O \rightarrow \text{Subj R/SO}$	5a. $\text{Obj} \rightarrow \text{N Acc}$
7a. $\text{Acc} \rightarrow \text{wo}$	14a. $R/SO \rightarrow \text{kaimasu}$
14b. $R/SO \rightarrow \text{kenkyuushimasu}$	14c. $R/SO \rightarrow \text{tabemasu}$

図3: 新しい規則1：日本語文法1に追加して新出文法項目を説明する

自由語順については、 $R/S \rightarrow \text{Obj R/SO}$ と $R \rightarrow \text{Subj R/S}$ を参考にして、他動詞 R/SO が、最初に主語を取って R/O となり、それが次に目的語を取って R となるというふたつの規則によって説明する(規則 3a、3c)。

図3中の規則を日本語文法1に加えた文法、「メモ帳」ファイル「JGScrmblng2」を構文解析器で語列/kodomo wa sakana wo kenkyuushimasu/や/sakana wo kodomo ga tabemasu/に関して試験しよう。

2. 3 説明力向上のヒント2：どの部門で説明するかを見極める

現象を統語・意味論で説明するのか、語用論で説明するのかを見極め、適切な部門の規則を使う。規則によっては、日常の言語使用では、それを破った例がよくあることを説明できるだろう。たとえば、「他動詞」、「対格」、「目的語」、「自由語順」を学んだら、(2)中のような文について学生から質問があるだろう。規則 $R \rightarrow \text{Topic R/S}$ に従って、主語の要求が満たされた他動詞 R/O が、話題 Topic を取って R となる規則が考えられる(規則 3d)。そうすると、統語・意味論上、他動詞 R/SO が話題を取って R/S や R/O になる規則を拒めない(規則 3e、3f)。学習者は実際にこのような質問をする。

3d. $R \rightarrow \text{Topic R/O}$	3e. $R/S \rightarrow \text{Topic R/SO}$
3f. $R/O \rightarrow \text{Topic R/SO}$	

図4: 新しい規則2：日本語文法1にさらに追加する

日本語文法1に図3と図4中の規則を加えた文法(メモ帳ファイル「JGScrmblng3」)は、たとえば、語列/sakana wo kodomo wa tabemasu/を次のようにR(文)だと予測し、図5のようにその統語構造を分析する。

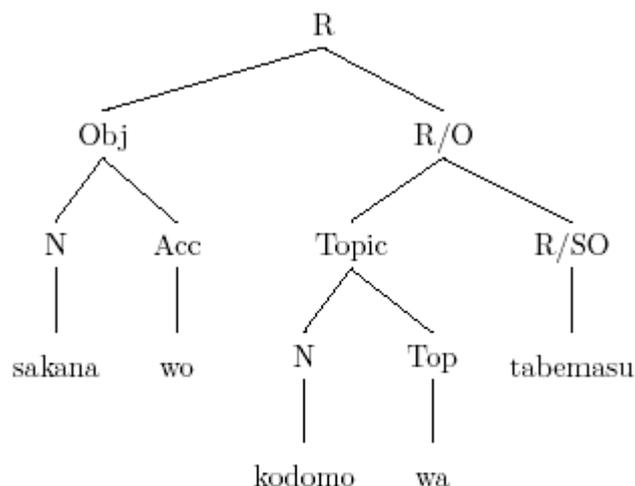


図 5: 新文法による 'sakana wo kodomo wa tabemasu'

語列 (2 a) 「さかなを こどもは たべます」は確かに変に聞こえる。だが、実は人々はよく使っている。ここで、Lee (2004)のように、語用論規則「話題は発話の初頭に続ける」によって説明すれば、語列 (2 a) から語列 (2 e) が変に聞こえるが、話題が対比と解釈され使われることがある事実を説明できる。極小主義 (MP)、主要辞駆動句構造文法 (HPSG)、形式意味論、語彙・機能文法 (LFG)、最適性 (Optimality) など、理論言語学は、日々、進歩しており、日本語教師も、これらの研究を知るなどして、自分の文法説明を日々、探求・研鑽し、分かりやすい説明を目指したい。

参照

- Lee, Hanjung. (2004). "Minimality in a lexicalist optimality theory", in Arthur, Stepanov, Gisbert Fanselow, and Ralf Vogel (eds.), *Minimality Effects in Syntax*, 241-288, Mouton de Gruyter: New York.
- Morgan, Jerry L. (2000). *EARLEY Parser*. University of Illinois at Urbana-Champaign.