

日本語における節境界解析時の語彙情報の影響

村岡 諭¹
(九州大学大学院 人文科学府)

坂本 勉
(九州大学大学院 人文科学研究院)

1. はじめに

本研究では初分析(First-pass analysis)において埋め込み節の始まりを解析するとき、すなわち左側節境界設定時に、格助詞の情報に加えて名詞句の情報が利用されているか否かを検証する。

2. 先行研究

従来の関係節解析の研究では、Mazuka & Itoh (1995)は主語のみを再分析する SR(Subject Reanalysis)文に比べ、主語と目的語の両方を再分析する SOR(Subject Object Reanalysis)文では再分析を行う回数や、解析木の構造の違いにより、解析の困難さが増大するという主張を行った。それに対して Hirose & Inoue (1998)は解析器が単文から複文への再分析を行う際に、関係節の主要部名詞の持つ意味情報を利用しており、再分析の困難さは主要部名詞句の持つ意味情報に左右されるという主張を行った。

上記の先行研究では共に埋め込み節主要部の入力と同時、もしくはそれ以降に左側節境界を設定すると考えられていたが、Miyamoto (2002)では埋め込み節主要部の入力以前に左側節境界を設定すると主張した。そして文を解釈するために新たな節が必要となった地点に左側節境界を設定するという Local Assignment of Clause Boundaries (以下 LACB)を提案した。LACBに基づき「おばあさんが年寄りを学生が」という名詞句の連続は次のように分析される。解析器は「おばあさんが」と「年寄りを」が入力されると単文であると解析を行う。次に「学生が」が入力されると、「NPが NPを NPが」という名詞句の連続が生じる。この名詞句の連続を項としてとりうる動詞は日本語に存在しない。そのため、解析器は新たな節が必要となったと判断し、「学生が」の左側に埋め込み節の左側節境界を設定する。Miyamoto (2002)では以上の解析過程において解析器が利用する最も重要な情報は格助詞の情報であると分析を行っている。

3. 実験

被験者：九州大学に所属する大学生、大学院生 37 名。

手順：玄海堂製「LingaLab」(<http://www7.plala.or.jp/genkaido>)を使用し、被験者ペースの読み(self-paced reading)で、画面中央に各句ごと呈示し、それぞれの文節の読み時間を計測した。

刺激：「NPが NPを」と「NPが NPを」の双方を項として選択できる動詞(ex:好きだ)を埋め込み節の動詞として用いた。刺激文は有生性(有生名詞句 vs.無生名詞句)と格助詞(が vs.を)による二要因二水準で構成されている。a-b, c-d 間は P5 における格助詞のみが異なり、a-c, b-d 間は P5 の名詞句の有生性のみが異なる。格助詞の要因に関する比較は被験者内で行い、有生性の要因に関する比較は被験者間で行った。

- | | | | | | | | | | | |
|-----|-------|------|-----|-----|-----|-----|-----|-----|------|-----|
| (1) | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| a. | 議員たちに | きさくな | 首相が | 話題の | 外相が | 好きだ | という | 事実が | 広まった | 。 |
| b. | 議員たちに | きさくな | 首相が | 話題の | 外相を | 好きだ | という | 事実が | 広まった | 。 |
| c. | 議員たちに | きさくな | 首相が | 話題の | 雑誌が | 好きだ | という | 事実が | 広まった | 。 |
| d. | 議員たちに | きさくな | 首相が | 話題の | 雑誌を | 好きだ | という | 事実が | 広まった | 。 |

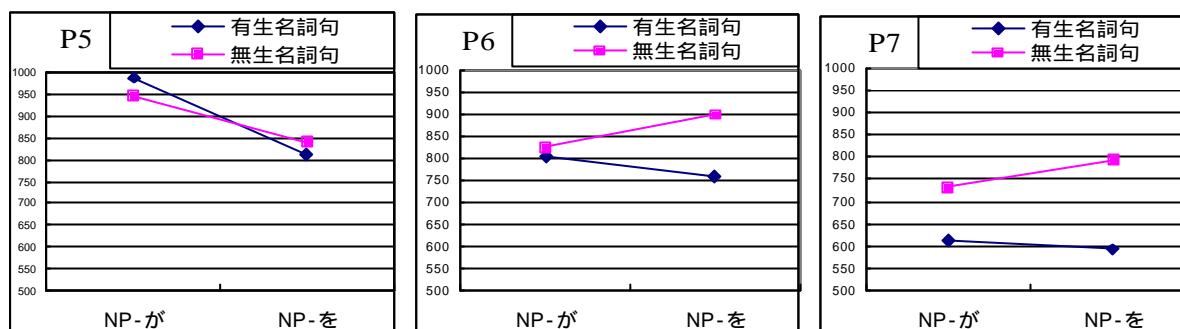
解析器に対する名詞句の語彙情報による影響が無いのであれば、有生性の要因による読み時間の変化は観察されないはずである。よって P6~P10 において(1a)と(1c)の読み時間に差は生じず、(1b)と(1c)の読み時間に差が生じないという結果が予想される。

結果：P5「外相がを・雑誌がを」において格助詞の効果がみられ、「NPが」の読み時間は「NPを」の読み時間に比べ有意に長かった ($F_1(1,35)=11.25, p<0.005; F_2(1,26)=40.02, p<0.0001$)。P6「好きだ」において有意な交互作用が見られ ($F_1(1,35)=3.79, p=0.0595; F_2(1,26)=11.91, p<0.005$)、 F_2 分析において格助詞「を」における有生性の単純主効果が見られた ($F_2(1,52)=21.83, p<0.0001$)。P7「という」

各句における読み時間 (単位は MS)

読み時間 (単位:MS)	P5		P6		P7	
	NP-が	NP-を	NP-が	NP-を	NP-が	NP-を
有生名詞句	986	810	805	757	612	589
無生名詞句	945	841	824	898	730	794

において有生性の効果が見られ、無生名詞句の読み時間は有生名詞句の読み時間に比べ有意に長かった($F_1(1,35)=3.65$, $p=.0644$; $F_2(1,26)=109.26$, $p<.0001$)。さらに P7「という」において有意な交互作用が見られ、($F_1(1,35)=13.16$, $p<.001$; $F_2(1,26)=10.19$, $p<.005$)、無生名詞句における格助詞の単純主効果が見られた($F_1(1,35)=14.32$, $p<.001$; $F_2(1,26)=11.11$, $p<.005$)。



考察：P5における格助詞の効果は Miyamoto (2002)の結果と矛盾しない結果であった。しかし、解析器が左側節境界を設定する際に最も利用する情報は格助詞の情報であるという Miyamoto (2002)の仮説では説明できない事実が観察された。つまり、Miyamoto (2002)の提案した LACB では P5 入力時まで(1b)と(1d)において共に同じ解析木を構築することを予測するため、P6,P7 における交互作用を説明できないのである。

そこで、解析器が左側節境界を設定する際に、名詞句の格助詞の情報に加えて名詞句の語彙情報を利用して解析を行っていると考えらる必要がある。ここで P5 入力時まで(1b)と(1d)において異なる解析木が構築されていたため P6 における読み時間の差が生じたと仮定する。この仮定に基づくと解析器は(1b)では P6 入力以前に左側節境界を設定していたが、(1d)では左側節境界を設定せず三項動詞の出現を予測していたために(1d)に比べ(1b)の読み時間が長くなった分析を行うことができる。もしこの仮定が正しいのであれば、Miyamoto (2002)の分析は間違っており、名詞句の語彙情報が重要な情報として利用されているという結論に至るのである。

5. まとめ

本研究の結果は、解析が行われる際に解析器が格助詞の情報だけでなく名詞句の語彙情報を利用しているという仮説を支持するものであった。しかし、本研究では解析器がいかなる方略に従い、そしてどのように名詞句の語彙情報を利用しているかを述べるには至らなかった。

解析器が名詞の語彙情報をどのように利用するかに関する方略として Sakamoto & Walenski (1998)によって提案された Theta-checking strategy がある。この方略は、文の主要部が入力される前に解析器が格助詞の情報を利用して名詞句に仮の意味役割(tentative theta-role)を付与し、その後主要部が入力された際に解析木の構造を確定させるというものである。この方略に従うと名詞句が入力された時点で解析器は名詞句の語彙情報を参照して「仮の意味役割」を付与することになり、名詞句の語彙情報を利用することになる。しかし、Sakamoto & Walenski (1998)には解析器が「どの格助詞でマークされた名詞句にいかなる仮の意味役割を付与するか」を明示的に示してはいないという問題点がある。

今後残された問題として日本語に存在する動詞がとることの出来る名詞句の連続においても、解析器は格助詞や名詞句の語彙情報を利用して左側節境界を設定しているのか否かを検証し、初分析における格助詞と仮の意味役割の関係を明らかにする必要がある。

参考文献

- Hirose, Y. & Inoue, A. (1998). "Ambiguity of reanalysis in parsing complex sentences in Japanese". In D. Hillert. (ed.). *Sentence processing: A crosslinguistic perspective: Syntax & Semantics, Vol. 31*. 113-147. San Diego, CA: Academic Press.
- Mazuka, R. & Itoh, K. (1995). "Can Japanese Speakers be Led Down to the Garden Path?" In R. Mazuka. & N. Nagai. (eds.). *Japanese Sentence Processing*. 295-329. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Miyamoto, T. E. (2002). "Case Markers as Clause Boundary Inducers in Japanese". *Journal of Psycholinguistic Research, Vol. 31*. 307-347
- Sakamoto, T. and Walenski, M. (1998). "The processing of empty subjects in English and Japanese." *Sentence processing: A crosslinguistic perspective: Syntax and Semantics, 31*, 95-111. San Diego, CA: Academic Press.

ⁱe-mail address: muraoka@lit.kyushu-u.ac.jp