

# 活性度に基づく文処理モデルと木接合文法による容認性予測\*

高橋 慶

東北大学大学院国際文化研究科

〒980-0751 宮城県仙台市青葉区川内 41

kei-ta@mail.tains.tohoku.ac.jp

## 1 背景

理論言語学において容認性を予測する際、構造性を仮定することにより様々な言語現象の説明に成功してきた。しかし、語順が容認性に影響を与えていると考えられる現象も存在し、それらに対しては構造性による分析のみでは不十分である。以下に例を示す。

- (1) a.??The elderly man told about [that the legend was wrong] and [the truth].  
b.?The elderly man told about [the truth] and [that the legend was wrong].
- (2) a.??The executive producer denied, but the general manager agreed with [that the project is interesting].  
b.?The general manager agreed with, but the executive producer denied [that the project is interesting].

(1)において、*that* 節は前置詞 *about* の項として現れることを認めていない。しかし、(1a-b)における対称性に見られるように、*conjunct* の順番の変化に応じて容認性に影響を与えている。また、(2)では *that* 節を項として許容する不完全動詞句と許容しない不完全動詞句の等位接続される順番によって容認性が変化している。以上から語順による容認性への影響が観察される。

また、以上の現象から観察されることとして、「headは項に対し「近い」際に統語的制約を課す」ということが挙げられる。これは Moosally (1999)、Sadock (1998) などにおいても指摘されていることである。しかし、定式化された統一的説明は行われていない。また、(3)のような、構造的に影響を与えないとされるような文においても容認性の差が見られる。

- (3) a.\*John was thinking about that he was wrong.  
b.??John was thinking about (pause) that he was wrong.  
c.?John was thinking about, by the way, that he was wrong.

\*本研究は一部、東北大学 21 世紀 COE プログラム「言語・認知総合科学戦略研究教育拠点」および東北大学学際科学国際高等研究センタープログラム研究からの研究経費の支援を受けて行われている。

容認性については、理論言語学においては、“\*”、“?”などの記号を使用することにより容認性の程度を表してきた。Bard et.al. (1996)、Keller (2003) などでは容認性は連続的であるとしている。本稿では後者の考えにならない、容認性は連続値によって表されると考える。<sup>1</sup>以上のことから、一連の現象は文法的、もしくは統語的問題ではなく、処理器の問題として捉えることにする。

## 2 処理モデル

本セクションでは Takahashi, Ishikawa and Yoshimoto (2006) において提唱されている、「活性度に基づく文処理モデル」を概観することにする。

### 2.1 活性度

本モデルにおいては、全てのノードは活性度を属性として持つと仮定している。以下にその概略図を示す。*av* は活性度カウンターとし、ノード名 (*X*, *Y*, *Z* など)、ノードラベル (NP, VP など) を引数とし、活性度は実数により表される。また、活性度の和は常に定数

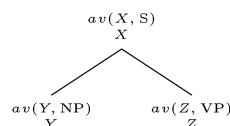


図 1: 活性度情報付きの “S → NP VP” の木構造

(*T*) であると仮定する。

$$(4) \quad T = \sum_{i=1}^n av(X, CAT_i)$$

本モデルでは、活性度の以下の条件に基づく非活性化により、統語情報が失われると仮定している。

- (条件 1) semantic predicate-argument structure が完成したと処理器が判断した時  
(条件 2) 処理時間経過、もしくは、overt phrase が挿入された時

<sup>1</sup>容認性が離散値であるか連続値であるかの議論については本稿では扱わないことにする。また、便宜上、本稿では容認性の指標は従来の言語学における指標に倣う事にする。

(条件 1) では、意味内容が得られ次第急激に統語情報は非活性化されることを意味する。これは、作業記憶の容量は有限であること、また、統語情報は意味内容を得るために利用されることを背景としている。これらの非活性化を以下減衰関数により定義する。

- (5) a.  $f =_{\text{def.}} \lambda X.\text{id}(X) \ 0$  (predicate-argument structure が完成したと仮定された際に適用される)
- b.  $g =_{\text{def.}} \lambda X.\text{id}(X) \ av(X, \text{CAT}_i) = av(X, \text{CAT}_i) - a$  (処理時間経過に伴い適用される)
- c.  $h =_{\text{def.}} \lambda X.\text{id}(X) \ av(X, \text{CAT}_i) = av(X, \text{CAT}_i) - b$  (overt phrase が挿入された時に適用される)

関数  $g$  において、非活性化は処理時間の経過に伴い繰り返し起こることから、関数  $g$  の適用回数は以下の式によって計算される。  $c$  は定数、  $t$  は処理経過時間を示す。

$$(6) \quad \text{cycle} =_{\text{def.}} \frac{t}{c}$$

よって、  $n$  サイクル後の活性度は (7) のように示される。  $av_0(X, \text{CAT}_i)$  は初期状態の活性度を示す。

$$(7) \quad av_n(X, \text{CAT}_i) = av_0(X, \text{CAT}_i) - na$$

処理中に曖昧性が生じる場合、本モデルでは、全ての処理の可能性を並列的に作業記憶に保持すると仮定している。しかし、活性度の高い順に並べられているために活性度の高い処理ほど優先される。処理器による  $i$  通りの予測されたカテゴリーの中に、実際の入力と同じカテゴリーが存在する場合、他の全ての可能性の持つ活性度は以下の式によりマッチした処理の活性度に集約される。

$$(8) \quad av(Z, \text{CAT}_k) = \sum_{i=1}^n av(Z, \text{CAT}_i)$$

一方、いかなる予測にもマッチしないカテゴリーが入力された場合、そのカテゴリーの活性度は以下の式により計算される。

$$(9) \quad av(Z, \text{CAT}_j) = 1 - \sum_{i=1}^n av(Z, \text{CAT}_i)$$

また、予測されたカテゴリーが非活性化されるに伴い、活性度の和は一定であることから、減衰した活性度が当初処理器によって予測されていなかったカテゴリーの活性度となる。  $m$  種類の予測されなかったカテゴリーが存在すると仮定すると、それぞれの予測されなかったカテゴリーの活性度は以下のように計算される。

$$(10) \quad av(Z, \text{CAT}_m) = \frac{1 - \sum_{i=1}^n av(Z, \text{CAT}_i)}{m}$$

## 2.2 活性度と容認性

本モデルでは、容認度は各ノードの活性度により実数値として予測される。その際、適格性関数 (well-formed functions) により計算される。  $Z_1 \dots Z_i$  を daughter とし  $Z$  を mother node とする local tree の適格性は以下のように計算される。

まず、自身のノードの適格性を関数  $\text{wf}_1$  により求める。(11) にその定義を示す。

$$(11) \quad \text{wf}_1(Z_i) =_{\text{def.}} av(Z, \text{CAT}_i)$$

次に、ノード  $Z$  の daughter ( $Z_i$ ) 全体の適格性を関数  $\text{wf}_2$  を定義することにより求める。

$$(12) \quad \text{wf}_2(Z) =_{\text{def.}} \prod \text{wf}_1(Z_i)$$

最後に mother node 自身の適格性 ( $\text{wf}_1(Z)$ ) と daughter の  $\text{wf}_2$  値を掛け合わせることで、local tree の適格性が得られる。

$$(13) \quad \text{wf}(Z) =_{\text{def.}} \text{wf}_1(Z) \times \text{wf}_2(Z)$$

以上の計算をトップノードの適格性が得られるまで繰り返すことにより、容認度として値が得られる。

## 3 木接合文法を用いた処理過程

前節では作業記憶内における、活性度をもとにした処理モデルを概観してきたが、文解析を行うにあたり、どのような統語情報がどのような論理処理に基づいて実現されるかについて議論する必要がある。本稿では Tree Adjoining Grammar (TAG; Joshi and Schabes 1992 等) に基づき、本モデルとの融合によりどのように文が解析されていくかを述べる。TAG が本モデルにおいて有用である理由として、まず、語彙化された文法であるために計算量が少なくすむという点が挙げられる。また、TAG では、語の持つ情報が構造化されており、本モデルにおいて仮定されている「先読み処理」についても実現されている。また、心的側面として、インクリメンタルな処理も可能であることが挙げられる。特に等位接続構文においては、標準的な combinatorial categorial grammar と心理実験により比較した結果、TAG の方がより実際の人間の等位接続構文の処理に近いと述べられている (Sturt and Lombardo 2005)。

TAG を採用するにあたり、(14) を仮定する。

(14) initial/aux tree の持つ活性度は、語の持つ活性度と等価である

仮定 (14) は、TAG においては語が長期記憶から呼び出される際、語彙構造自体がノードを内在する木構造として表現されるため、語の構造内のノードは全て同じ活性度を持つことを意味する。以下に initial tree である *bold* を例として示す。<sup>2</sup> また、TAG 上の操作である adjunction、substitution について以下を仮定する。

(15) adjunction、substitution 操作の際には、操作されるノードの活性度のみ置き換えられる

<sup>2</sup>以下では各ノードの活性度を単に [activation value] として表す

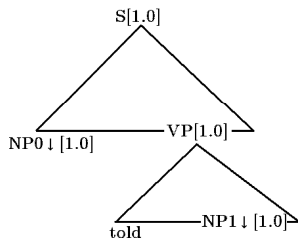
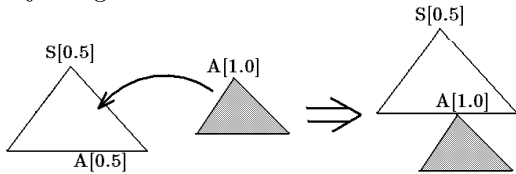


図 2: *told* の活性化度付きの語彙構造

TAG (特に Feature-based TAG) などでは各ノードは top feature、bottom feature を持ち、自身のノードの下にどのような語彙情報を持つものが接合出来るか、もしくは接合出来ないかについて制約が記述されている。項が入力されると、mother node の bottom feature と入力 の top feature が単一化されるが、活性化度については項の入力により未完成句が具現化されるという点から入力の活性化度が override されると仮定する。substitution 操作についても同様に、置き換えられたノードの活性化度は override される。

- adjoining



- substitution

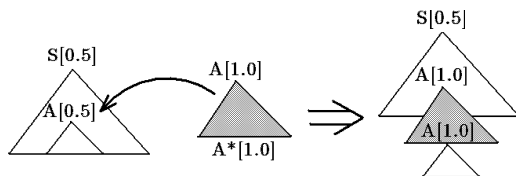


図 3: 活性化度付きの adjoining、substitution 操作

また、容認度を得るための適格度計算について、TAG では図 4 のように initial/aux tree が構造をなしているために、適格性は、各 initial/aux tree 単位で行われる。

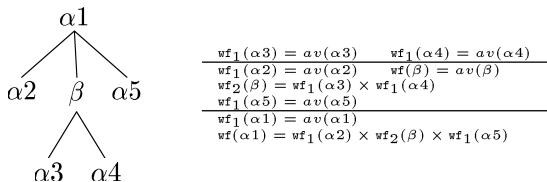


図 4: TAG 上での適格度計算

本モデルと TAG により等位接続構文 (1b) は以下の

ように処理される。処理手順を図 5 に示す。<sup>3</sup> まず、aux tree である *the*、initial tree である *elderly*、*man* が順に長期記憶から呼び出され、名詞句を形成する。本モデルでは「先読み処理」を仮定しているので、*the* が呼び出された時点で、具現化されていないトップノード S が形成される。

次に動詞である *told* が入力され、項として図 5 に示されているように、処理候補  $\beta 2$  と  $\alpha 2$  が現れ、並列的に処理されて活性化度順に作業記憶に保持される。もし最も活性化度の高い parse が実現しない場合、活性化度の再計算により計算量が増え、それが処理遅延時間として反映される。この例では S を項に持つ  $\beta 2$  の活性化度が最も高いが、入力は名詞句 ( $\alpha 3$ ) であるために活性化度は計算式 (8) により  $\alpha 2$  をもとにした処理候補に集約され、 $\beta 2$  の処理は棄却される。<sup>4</sup>

第 1 conjunct である名詞句  $\alpha 3$  が入力されると、この時点で文全体の容認度  $wf(S)$  が計算式 (11)-(13) によって計算され、predicate-argument structure が完成するために減衰関数  $f$  により各ノードの活性化度は急激に減衰する。よって、各ノード自体の情報、及び、各 initial/aux tree 内の情報は残るが、語同士の依存関係情報は失われる。よって、 $\alpha 2$  が課すべき統語制約は失われるために、*and* が入力され、第 2 conjunct が入力される際には統語情報を元にせず modus ponens が行われ、結果として意味的不適格がないため *that* 節を許容する。<sup>5</sup>

Right Node Raising 構文 (2) の場合においては、第 2 conjunct が処理される間に、減衰関数  $g$  が繰り返し適用され、減衰関数  $h$  もまた適用されるために第 1 conjunct 内の前置詞の活性化度は十分に非活性化され、bottom feature の情報が失われるために、第 2 conjunct 内の前置詞の要求に従い、*that* 節を容認できると分析できる。

## 4 考察

以上では、活性化に基づく処理モデルが TAG 上でどのように処理されるかについて分析してきたが、心理言語学的に知られているガーデンパス効果、及び、syntactic priming 効果についても以下のように分析することが可能である。

### 4.1 ガーデンパス効果

ガーデンパス効果とは、予測される文処理が失敗を起こし、再分析などを要することにより処理遅延が見られる現象である。本モデルでは、活性化を仮定することによって、その差からガーデンパス効果を予測することが可能である。

(16) The horse raced passed the barn fell.

<sup>3</sup> 図 5 の 1 列目には入力後の処理可能性 (candidate)、2 列目には実際に入力を示す。第 1 行目の candidate 欄には、NP *the elderly man* を処理した後の処理可能性を示す。

<sup>4</sup> 各処理候補の活性化度の合計は常に 1.0 であることに注意されたい。

<sup>5</sup> 図 5、5 行目の candidate は、意味的に不適格のない候補を示す。統語的に予測される処理は存在しない。

candidates	input	operation	output( semantics )	deactivation ( funct : value( node ) )	wf(S)
$\beta_1$ : VP[0.5] $\alpha_1$ : S[0.5] 	told[1.0]	adjunction	$\lambda x(z, \text{the elderly man}') \\ \lambda \text{told}'(z, \text{the elderly man}')$	N/A	N/A
$\beta_2$ : S[0.6] $\alpha_2$ : PP[0.4] 	about[1.0]	adjunction	$\lambda \text{told about}'(z, \text{the...}')$	$g:1.0-0.1=0.9(\alpha_1)$ $g:1.0-0.1=0.9(\beta_1)$	N/A
$\alpha_3$ : NP[1.0] 	the truth[1.0]	adjunction	$\lambda \text{told about}'(\text{the truth}, \text{the...}')$	$f:0(\alpha_1)$ $f:0(\beta_1)$ $f:0(\beta_2)$	1.0.
none	and[1.0]( $\beta_3$ )	substitution	$\lambda \text{told about}'(\text{and}'(\text{the...}, z), \text{the...}')$	N/A	N/A
$\beta_4$ : S[0.5] $\alpha_4$ : NP[0.5] 	that ...[0.5]	adjunction	$\lambda \text{told about}'(\text{and}'(\text{the...}, \text{that...}, \text{the...}')$	$g:1.0-0.1=0.9(\beta_3)$	0.5

図 5: (1b) の主語 NP 以降の処理過程.

(16) において、主語 NP *the horse* を処理した時点で処理器は動詞句と関係詞節を予測する。前者の活性度を 0.8 とし、後者の活性度を 0.2 と仮定すると、*barn* の処理後、減数関数  $f$ ,  $g$  によって十分に非活性化されているが *fell* が入力されることにより、関係詞節が入力されるといふ予測の活性度を最優先にする必要がある。その際、動詞句予測の分の活性度と、処理時間経過による減衰によって生じたその他の予測可能性の活性度を再計算する必要があり、計算量が増え、作業記憶に負担がかかる。よって活性度の再計算にかかる負担が処理時間の遅延として表われると分析できる。

## 4.2 syntactic priming

syntactic priming とは、一度処理された構文が再度提示されると処理が促進されるという現象である。本モデルでは、TAG を採用することによりこの効果を予測することが出来る。

(17) I just want to give [a present] [to my son].  
 Ummm... but I don't know how to give [it] [him].

TAG においては、例えば *give* のような transitive、di-transitive の 2 つの構文をとる動詞においては個別に長期記憶に保存されていると仮定されている。よって、一度呼び出された項目は他方に比べ活性化していると考えられる。よって、同じ動詞で再度同じ構造を持つ項目が表われた場合、より高い活性度を持つ項目の方がより処理が促進されると考えられる。

## 5 まとめ

本稿では活性度に基づく文処理を概観し、語彙化された文法理論である TAG を適用することによって、文法のみでは分析困難であった言語現象に対し、統一的な説明を与えることに成功した。しかし、まだ心理実験による容認性の緻密な検証が行われていないことが問題点として挙げられる。そこで、線形的な容認性判断とともに、回帰分析による減衰関数における減衰値の設定が優先的課題として挙げられる。また、本モデルは作業記憶内での処理を仮定している。先行研究としては、Baddeley (2003) などが挙げられるが、言語特有の解析

済み領域は特定されていない。よって、作業記憶モデルへの適用も今後の課題として挙げられる。

## 参考文献

- Baddeley, Alan. 2003. "Working Memory: Looking Back and Looking Forward". *Neuroscience* October, Vol. 4. 829–839.
- Bresnan, Joan. 2000. *Lexical-Functional Syntax*. Oxford: Blackwell.
- Bard, Ellen G., Dan Robertson and Antonella Sorace. 1996. Magnitude Estimation of Linguistic *acceptability*. *Language* 72, No.1, 32–68.
- Falk, Yehuda N. 2001. *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford: CSLI Publications.
- Joshi, Aravind, and Yves Schabes. 1992. Tree-Adjoining Grammar and Lexicalized grammars. *Definability and Recognizability of Set of Trees*, eds. M. Nivat and M. Podelski. 31–81. Princeton: Elsevier.
- Keller, Frank. 1996. Gradience of Grammaticality: Experimental and Computational Aspects of Degrees of Grammaticality. Doctoral Dissertation, University of Edinburgh.
- Moosally, Michelle J. 1999. Subject and Object Coordination in Ndebele: An HPSG Analysis. *WCCFL 18*, eds. S.Bird, A Cammie, J. Haugen, P. Norquest, 379–392. MA: Cascadilla Press.
- Sadock, Jerrold M. 1998. Grammatical Tension. *CLS 34: The Panels*, 179–198.
- Sturt, Patrick, and Vincenzo Lombardo. 2005. Processing Coordinated Structures: Incrementality and Connectedness. *Cognitive Science* 29, 291–305.
- Takahashi, Kei, and Kiyoshi Ishikawa and Kei Yoshimoto. 2006. An Activation-based Sentence Processing Model. *Proceedings of The 20th Pacific Asia Conference on Language, Information and Computation*. eds. Tingting He, Maosong Sun and Qunxiu Chen. 303–310. Beijing: Tsinghua University Press.