

# 改心の理由

私が (突然) 意味タグづけの研究なんかを始めたわけ

黒田 航

独立行政法人 情報通信研究機構 知識創成コミュニケーション研究センター

Modified on 10/15/2007, 02/01/2010; Created on MM/DD/2005

## 1 はじめに

私は博士号取得後の二年ほどの留学<sup>1)</sup>を終え、日本に戻って来て、次の二つを心に決めた:

- (1) 真剣に意味の研究をしよう
- (2) 真剣に日本語の研究をしよう

私はこれを一種の「改心」と見なす。この心境の変化の背景になった経験を説明するのがこのエッセイの目的である。

## 2 私が意味論を(真面目に)しなかった(言い)わけ

### 2.1 私が意味論から遠ざかっていたわけ

私は大学院生時代、あまり真剣に意味論、語用論の研究に取り組まなかった。これは私の出身研究室がどこであるかを考えると、少し意外かも知れない。しかし、別にこの分野が苦手だったわけでも、嫌いだったわけでもない。理由は主に二つある。

- (3) すでにデキ上がった土台で研究をするのは面白くない

<sup>1)</sup>私はUCSDのCenter for Research in Language (CRL) [<http://crl.ucsd.edu/>]という組織に二年弱の期間に渡って留学した。CRLはSimple Recurrent Networkで有名なJeffrey Elman [<http://crl.ucsd.edu/~elman/>]がいるところである。そこで私は今は亡き偉大な言語心理学者Elizabeth Bates [<http://crl.ucsd.edu/~bates/>]にも大いに世話になり、CRLでは、Jeffの指導するドイツからの留学生を含め、認知科学科の大学院生との交流を含めて、非常に有益な経験をしたと思う。この辺の事情に関しては、私の回想録[40]を読んで頂きたい。

一つだけ補足的に言っておくと、日本の大学にはUCSDのCognitive Science Department (ここでは「認知科学科」と呼んだ)に相当する学科はないと思う。工学部の電気計算機科とロボット工学科に、理学部の医学部の大脳生理学科が継ぎ合わされ、それに知覚心理学科と言語学の一部が混じり交じったような、又工のようでありながら、非常に理工系的色彩の濃い学科だった。

- (4) 誰かがマジメに認知科学的な観点から統語論を研究するための基盤—できればコネクショニスト模型などと互換性ある統語論の基礎—を築かなければ

それぞれについて以下に簡単に説明しておこう。

### 2.2 デキ上がったものはつまらない

第一の理由は、私はいわゆる「認知(言語学)」系の枠組み、例えばLakoffの認知意味論(Cognitive Semantics) [26], Langackerの認知文法(Cognitive Grammar) [28, 29], Fauconnierのメンタルスペース理論(Mental Space Theory) [6], ないしはブレンド理論(Blending Theory) [7, 8, 9, 10]の「枠組み」で意味論研究をする気がなかったからである。私はモノを作るのが好きな方なので、「すでにデキあがっているものを使う」のは好みではないのである<sup>2)</sup>。

更に言うと、私はイメージ先行の意味論には興味がなく、院生当時はCategorial Grammar (CG), Generalized Phrase Structure Grammar (GPSG), Montague Grammar (MG), Tree Adjoining Grammar (TAG)などに関心をもっていた<sup>3)</sup>。イメージで意味を記述するなんて、よっぽどうまくやらないと失敗するに決まっていると思っていたし、今でも思っている。それは

<sup>2)</sup>とは言え、これは認知意味論、認知文法、メンタルスペース理論/ブレンド理論が仮にデキあがっている、とすればの話だが。

<sup>3)</sup>今日の知見で見ても、これらのモデルは優れている。それは統語構造と意味構造をア prioriに切り離さない、という思想に基づいているからである。ただ、方法的に意味構造の記述を統語構造の記述から分離できるし、言語理論のインフラ構造を確立するためにはその必要があるというのは私の現在の見解であるけれど、言語の本質は確かに意味構造と統語構造の不思議な結びつきなのである。重要なのは、例えば[28, 29]が「文法はシンボリックだから」とか何とか言って、結局は統語構造の記述に「手抜き」をしているような反則はイケナイということである。これはチョムスキー派の言語学者が「意味構造と統語構造が別のもんだから」とア prioriに理由をつけて切り離すのと、やっていることは実質的に変わらない。

不可能なことではないが、ヒトは視覚優位な生物なので、図にちゃんと記述されていないことでも、読み手が勝手に情報を補正して、書かれていないことまでも読み取ってしまう可能性が高い。こうなると、図像化によって本当に記述されていることと図像化の解釈の際に(知らない間に)補われていることの区別がつかず、記述/説明の実質的效果が明らかではない。この補正の効果を含めて「説明」と称するのであれば、それは錬金術と何ら変わるところはない。いずれはそこから「意味の化学」が生まれるかも知れないが、そのときまで私は生きていないだろう。

### 2.3 認知言語学の枠組みの範囲でも、誰かが真面目に統語論をしなければ...

第二の理由はこれより深刻で、決定的なものである。

私は当時、意味論をやらなければならない理由をそれほど強く感じておらず、それ以上に統語論を真面目にやらなければならないと感じていた。なぜか?

第一に—これは後にどうも誤りだったと判明したが—私がやらなくても、認知言語学の分析手法は自然に進歩するだろうと楽観していた<sup>4)</sup>。

こちらが本質的な理由だが、第二に、私は認知言語学の偉大な指導者がそうしているように、統語論を真面目にやらないで生成言語学を批判するのは、知的誠実さの点でガマンがならなかった<sup>5)</sup>。だから私は、彼らが真面目に統語論をやらないことを正当化するような研究プログラムに従うことを潔しとは

<sup>4)</sup>裏話をすると、当時から私は自分の指導教官 Y 教授と認知言語学の「お絵かき」の評価を巡って大きく意見が食い違っていた。だが当時は、Y 教授の「今はまだ過渡期だから!! 暖かい目で見守ってやってくれ!!!」という説得はもっともらしく聞え、効果があった。実際、私はそれに宥められ、しぶしぶ口を閉ざした。だが、10 年経って見て、何も事態は変わっていない、いや、むしろ生成言語学と認知言語学が完全に分離することによって緊張感がなくなり、事態は悪化しているのではないかととも思われる。この辺が、私が最近、認知意味論の基礎づけ、認知文法の基礎づけに対し、批判的な言説を公にするようになった理由の一つである。

<sup>5)</sup>Lakoff, Langacker, Fauconnier はいずれも、確かにかつては生成文法の統語論をやっていた。だが、認知言語学に加担するようになって、彼らのやり始めたことは、統語論と意味論との関係を公平に結びつけることではなく、極限まで拡大された意味論に統語論を還元することだった。それは、生成意味論が、極限まで拡大された統語論に意味論を還元しようとして失敗したのと、正対的なアプローチである。これはまったくバカらしい試みで、失敗するのは目に見えている—少なくとも私にはそう思えるのだ。これらの運動の主導者—特に Lakoff—は、どうやら複雑な世界には耐え切れないうまく、世の中が真っ白か真っ黒かのどっちかではないとダメなようだ。

せず、生成言語学以上に真剣に統語論を、根本的なレベルからやり直す仕事を手がけた。

特に重要だったのは、Elman [5] などのコネクショニズムの研究成果と“実質的”な互換性をもつ記述の枠組みを作ることだった<sup>6)</sup>。この仕事の内容に関しては、指導教官と少なからずイザコザがあったが、これも今となっては懐かしい思い出だ。このような難儀な仕事も博士論文の完成 [25] と共に一段落した。

ただ、博士論文が完成した頃、私は率直に言って言語学に飽きていた—正確にはその非科学性に辟易していた—ので、まったく別のことをやってみることにした。これが UCSD の CRL に留学したのきっかけである<sup>7)</sup>。それを機会に私は言語学のことはきれいさっぱり忘れて、以前から関心をもっていたコネクショニスト模型 (Connectionist Models) の勉強をすることにしたわけである<sup>8)</sup>。

### 2.4 CRL での衝撃的な経験

私はこういう経緯で CRL で UCSD 内外の院生に交じってコネクショニスト研究に手を染めたが、その際、Penn TreeBank [1] で遊ぶ機会を得た。これが私が言語資源の重要性を認識する初めての機会だった。

私は Jeff がコネクショニスト模型を使った研究から、コーパスに研究に興味を移しているのを知った。彼は M. Tomasello などと連絡を取り、CHILDES<sup>9)</sup> コーパスの解析で博士号を取ろうとしてる Holger Keibel という名の院生を指導していた。意外なこと

<sup>6)</sup>私がここで実質的という点を強調するのは、Langacker [29, pp. 525–536] の互換性の主張は、単に用語上の互換性 (terminological compatibility) である可能性が高いからだ。

私がこう言うのには、それなりの理由がある。Langacker はコネクショニスト模型で認知文法をモデル化したいと言う UCSD の認知科学科の院生の申し出を「そんなことしなくても、私の理論は正しいから」と言って退けたことがあるからである。私はこの話を、その院生本人から聞いた。

<sup>7)</sup>私は UCSD に Fauconnier, Langacker がいるのは知っていたが、彼らの仕事に特別な関心はなかった。Fauconnier の講義には少し関心があったのだが、彼は Dean の仕事で忙しく、私が滞在中に講義はなかった。これは残念と言えば残念だ。

<sup>8)</sup>「以前から関心をもっていた...」というのは、非常に控え目な表現である。私は博士課程の三年目のある日、たまたま Elman [5] を読んで感激し、完成間近だった博士論文を放棄し、ゼロから PMA の方向へ書き直した。

<sup>9)</sup>今は亡き Liz Bates と Brian MacWhinney が共同して立ち上げた言語獲得研究プロジェクトで、チョムスキー派とは異なり、言語獲得の論理的問題 (Logical Problem of Language Acquisition) を単なる思弁に訴えず、可能な限り実証的な手段によって解決することを、その旨とする研究プロジェクトである。詳細に関しては、<http://childes.psy.cmu.edu/> を参照。最近は部分的に日本語化されてもいるらしい: [http://www013.upp.so-net.ne.jp/lang-ac/childes\\_databaseinfo.htm](http://www013.upp.so-net.ne.jp/lang-ac/childes_databaseinfo.htm)

だったが, Jeff は言語獲得のミニチュアを接続主義者モデルによって行う研究の限界を強調した. “starting small” は言語獲得/習得の際の重要な境界条件の一つだが, 境界条件の全体像は接続主義者モデルから自ずと決まってくるわけではなく, それを言語獲得コーパスを使って見積もる必要があることを訴えていた. Tomasello との交流も, Holger の指導もその延長だったようだ.

Holger は私の同室だった. 彼は来る日も来る日も CHILDES コーパスから情報抽出をするための Perl<sup>10)</sup> スクリプトを編集し, それを統計的に解析するという仕事を続けていた. 私は Perl 使いではなく Python<sup>11)</sup> 使い — 正確にはそうなりつつある見習いの存在 — だったが, 私たちは同室だったので, コーパス処理の問題 — 特にパターンマッチングのための正規表現 (regular expressions) の工夫の仕方 — に関して, ときどき意見交換した.

CHILDES は興味深いものだったが, それには (少なくとも私の留学当時は) 日本語の習得データはなかった. だから, 英語などの外国語の習得に特別の関心のない私が CHILDES のコーパスを使ってやれることはなかった. そこで私は CHILDES には首を突っ込まず, Penn TreeBank [1, 30] で遊んでみることにした. それは当時私が取り組んでいた接続主義者モデルを使った形態素区切りの学習シミュレーションの訓練コーパス作成のための下調べという意味もあった.

## 2.5 Penn TreeBank の衝撃

Penn TreeBank に初めて接したときには正直に言って驚いたが, 最初の衝撃が収まるにつれ, 次のような, お互いに関連した二つのことを実感した:

- (5) 日本語にも Penn TreeBank みたいな解析済みコーパスがあったら, どんなにスゴイことだろう
- (6) Penn TreeBank は確かにスゴイ — 特に統語論の見地から見るとスバラシイ — が, ほかの目的 (例えば意味論) には意外と使えない...

これらについて順番に触れることにする.

## 2.6 日本語の統語解析済みコーパスがあったら...

まず (5) の問題について私見を述べよう.

正直言って, 日本語に TreeBank に相当する解析済みコーパスがないのは衝撃だった. 本当は「京大コーパス」<sup>12)</sup> という形で存在していたのだが, 私は当時, その存在を知らなかった — 正確には, 名前しか聞いたことしかなく, その実態を知らなかった.

これは私の出身が京都大学であることを考えると, 少し意外なことである. 京大コーパスを開発した長尾研究室は京大にあり, 私の元指導教官は他の研究グループと較べると繋がりはあるほうだと思っていたが, 実際はそうではなかった, ということなのだろう. やはり, 文系/工系系の断絶は, 見えない形で存在する.

### 2.6.1 日本語のコーパス言語学の立ち後れ

Penn TreeBank に触発されているいろいろ調べてみると, 次のことが解った:

- (7) a. Penn TreeBank は統語解析の付加情報つきで公開された最初のコーパスであるが, コーパスの代表格というわけではなく, コーパス言語学はアメリカではまだ発展途上であり, どちらかというといギリスで進んでいること (さすが経験主義の国だ)
- b. 日本語のコーパス言語学は立ち後れている (例えば, 日本語には今だにバランスコーパスがない) こと
- c. だが, 概して言うと, マルチバイトの問題もあって, Roman 文字セットを使わない言語圏 (いわゆる CJK 言語圏) のコーパス言語学は遅れていること
- d. コーパス言語学はどういうわけか, UNIX の世界と繋がりが深いこと

などである

<sup>10)</sup><http://www.perl.org/>

<sup>11)</sup><http://www.python.org/>

<sup>12)</sup><http://www.kc.t.u-tokyo.ac.jp/nl-resource/corpus.html> (cf. [50])

## 2.6.2 日本語の「使える」統語理論は不在である

(7b)の最大の問題は、日本語の統語論が充実していないことによるものだと私には思われた。日本語の統語構造がツリーでうまく書けないのは—いわゆる言語学者さんたちが扱う「オモチャ文」以外の現実の文章の解析をやったことのある人には—周知のことである。つまり、日本語のマトモな統語理論は存在しないのである<sup>13)</sup>。

と同時に、私はつくづく実感した: 日本語の言語学をちゃんとやらなければならないのだなあ...。私は別に英語が好きでも何でも無いが、チョムスキー学派の言語学の向こうを張ろうと思ったら、まずは英語で研究しないとイケないだろうと思ひ、その路線でしばらく研究を続けていた。だが、この段階で、俄然、自分の母国語である日本語の研究が立ち後れているのを見て、啞然としたわけである。

私は別に愛国精神があるわけでも何でも無いが、私は—生成言語学に限らず—いわゆる「理論」言語学の研究者が英語の分析では数量詞移動とかメトニミーとか微妙な問題をいろいろ議論しておきながら、その反面、ろくすっぽ日本語文の分析ができないのを知っており、それには前から呆れていた<sup>14)</sup>。まず、大抵の「自称」言語学者は、日本語文を適当な形態素に区切れないのである<sup>15)</sup>。言語学者が母国語の形態素区切りを、工学系の研究者が作った形態素解析プログラムに依存するようでは、まさに本末転倒なのであるが、実態はこれに限りなく近い。

<sup>13)</sup>今は、日本語のための HPSG 解析, LFG 解析 [51] などが充実してきてはいる。

<sup>14)</sup>この傾向は、今なお顕著である。例えば、岩波から出版されている『言語の科学』シリーズで、あれほど多くの著者の取り上げる例文が英語なのだろう? これはハッキリ言って、異常である。

<sup>15)</sup>これは言語学者の大半が、形態素の認定が文脈自由 (context-free) だと無反省に想定してきた結果である。意味を考慮に入れない形態素の認定は文脈自由でよいかも知れないが、形態素の定義が意味に言及する以上、これは自己矛盾である。形態素の認定の文脈依存性 (context-dependency) は、例えば英語のような言語では軽微な現象として無視して構わないかも知れないが、日本語のような言語ではそうではない。これは [?, ?] に詳しく論じられおり、形態素分割に対する対案が提案されている。

このような問題意識で例えば [49] を読むと、言語学者は形態素への分割がまるで“自明”な課題であるかのような想定をしているという印象は免れないし、実際に学会で支配的な態度はそうである。なお、自然言語処理の従来形態素解析へのアプローチの概観は、[46] などを。形態素と辞書との関係については、[38] などを参照するとよい。

## 2.6.3 「罪」の自白

これを白状するのは些か恥ずかしいが、隠さずに言っておこう: 私は一般言語理論と英語の理論には少なからず係わってきたが、日本語の言語学にはあまり真剣に係わってこなかった。理由は二つある:

(8) 第一に、日本語の統語解析が英語の統語解析などより数倍は難しいのが自分の経験から分かっていたから

(9) 第二に、とにかく日本語関係の文献を読むのが苦痛だったから<sup>16)</sup>

である。

第二の点に関しては、最後の §4 で改めて論じるが、補足的に一言しておく、私が「京大コーパス」の存在を知らず、それが KNP (Kurohashi-Nagao Parser) という名のプログラム [50] で“係り受け解析”<sup>17)</sup>されていたことも知らなかったのは、この辺にも理由がある。

## 2.7 意味の実証的研究を可能とするインフラが必要

TreeBank との出会いは衝撃的で一目惚れに近い感じだったが、時間と共に熱も冷め、次第に粗も見えるようになってきた。

今は改訂されているが、初版の TreeBank では文法関係が明示的にエンコードされていなかった。Version 2 以降は { SNP, ONP, ... } のような区別がアノテーションとして付加されるようになったが、初版では、複雑な検索パターンを使って主語名詞句を文から採りだすのですら、ものすごく面倒だった<sup>18)</sup>。これは、私が思うに「文法関係はツリー構造から復元できる」

<sup>16)</sup>私は日本人で、日本語を母国語とする人間だが、典型的な日本の文系の研究者が書いた論文、研究書を読むほど苦痛なことではない—研究の意図、必要性が不明、研究の手続きが不明、研究結果の評価基準が不明—もう何から何まで意味不明なのである。高校生の書いた書評にすらもっとマシなものがあるのではないかと私は常々疑問に思う。

<sup>17)</sup>“係り受け解析”は句構造文法 (phrase structure grammar) に基づく句構造解析 (phrase structure parsing) とは異なり、依存文法 (dependency grammar) に基づく依存解析 (dependency parsing) の一種である。依存解析はヨーロッパ言語の解析でよく使われている。形態論が豊かで語順が自由な言語は句構造解析ができないことが多いのがその理由である。

<sup>18)</sup>tgrep というツールは、grep の拡張版でツリーに対して正規表現で検索を可能にするツールなのだが、検索パターンを作るのが面倒だった。今は tgrep2 になっているようだ。

という当時の言語学会を支配していた理論的バイアスを反映したものだと思う。

これは要するに、英語の場合ですら、言語の意味の実証的研究を可能にするインフラが整っていないということである。問題は、意味にアプローチしようと思ったら途端、従来の手法では手詰まりになるということだ。

### 2.7.1 意味タグがついていたら...

言語学者のコーパス利用の基本はトークン (tokens) を収集し、数を数えることである。だが、何のトークンを? タイプ (types) なしのトークンはありません。だが、意味が対象の場合、タイプとは何か?

簡単に言うと、意味タイプ (semantic types) が意味タグ (semantic tags) という形で明示されていない限り、意味トークンは数えられないし、分布 (distribution) も調べられない。意味が問題である場合、品詞を数えてもしょうがない。品詞は、非常に間接的な仕方では意味には係わっていない。

同様に、統語解析も、思ったほど—少なくとも生成言語学が成立して以来、理論言語学者が「教義」を信じて期待するに至ったほど—は使えない。解析された統語構造は意味的実体 (entities) をエンコードしていないし—言語学の論文で例文に上がるような単文でない限り—意味と構造との相関すら認められない。つまり、構文効果を見つける場合ですら、意味のパターンがそれとしてエンコードされていないと、見つけるのに—苦労なのである<sup>19)</sup>。

この経験を通じて私が実感したのは、意味が調査対象である場合、付加情報 (例えばタグ) がついていないと、コーパスの利用価値は非常に限定されるということだった。なのに、大部分の研究者の関心は言語の意味にあるのだ。

これが私が「ああ、英語の統語論のことばかりやっていてもダメなんだ。意味のことも真面目にやらないとダメなんだ...」と実感した理由である<sup>20)</sup>。

<sup>19)</sup>この点に関しては、池原らの日本語の非線型表現のデータベース [39] が公開されることで、多くの言語学者が陥っている無知蒙昧状態に対する啓蒙に大いに貢献するだろうと期待する。

<sup>20)</sup>ただし、PMA で英語の解析を最優先したのは、あくまでも戦略的な決定の結果である。日本語の統語構造を PMA で解析することは、言語学会に出回っている言語理論に較べたら圧倒的にうまく行くことはわかっていた—少なくとも HPSG と同じくらいうまくいくだろう—ことはわかっていたが、些か厄介なゼロ承応に関する技術的な問題があり、それを決定する根拠が不十分であることを理由に、後回しした面が強い。そうでなければ、

例えば、Penn TreeBank に意味タグがついていれば、どんなに事情が違っていたことだろう。だが、当時は PropBank [16, 23, 24] はなかったし、WordNet [11] の定義を付加したコーパスが存在するらしいのを知ったのは、それからしばらくしてからのことだった。だが、これにしたって、私が有用だと考える意味タグづけとは違う。そして、これこそが私を日本語のための意味役割タグづけの研究 [45, 44] に向わせる最大の動機となったものである。

## 3 意味論への私見

### 3.1 遠回りの効用

以上のような理由から、私は意味論を本格的に始めるまで、ずいぶん遠回りした。だが、結果的には、統語論の経路を通過してから意味論に戻ってきたのは、自分にとって非常に幸運だったように思う。迂回の原因が何であれ、そうしていなければおそらく、他の多くの意味論の研究者と同じく、意味がいったい何であるかに関して、救いようなく混乱し、今だに頭を抱えて苦悶していただろう。統語論をしっかり学ぶこと—それに幻惑されず、しっかり一定の距離を取ってそれを習得すること—は、意味を観察するのに非常に有効である。

意味はなかなか「見えない」対象である。それが見えるようになるには、十分に修業をつまなければならぬ。統語現象の集中的な研究はそれを可能にする—統語理論に「魂」を売り渡さなければ。

### 3.2 “(統語) 変形” を意味記述の手段と見なすことの有効性

私が生成言語学で実践されている解析法で、非常にすぐれており、分野に抛らず有効だと思う手法が二つある:

- (10) a. 変形: 任意の要素に“移動”, “削除”, “付加”, “代入”などの操作を施し, その結果を(有意味性の保存, あるいは内容の変化の観点から)考察するという手法

私の博士論文 [25] はもっと多くの日本語の例文を扱っていたらと思う。

- b. (非)容認性の分析: 自然に存在する表現ばかりでなく、\*のつくような“容認不可能”な表現も考察することで、言語表現の有意味性を、非無意味、あるいは下有意味性との相対的關係で理解するという手法
- (11) あたり前の文の意味を考えるだけでは、どの意味がどの形態素から発生するかに関する直観、あるいは、ある種の意味はどの個別の形態素からも発生していないことに関する直観が得られにくい

これらをうまく組み合わせることで、有意味性の複雑な基盤を明示的に、かつ効果的に解析できるのは明白である。

実際、私は非生成系の研究者—特に認知言語学系の研究者—がこのような手法を(めったに)使わない理由がまったく理解できない。私はいわゆる“移動”や“削除”や“付加”や“代入”が生成文法家が言う意味で“文法現象”だとはゼンゼン思わないし、(容認可能性ではなく)“文法性”という概念が—理論上はともかく—実証的に意味のある概念だとは思わないが、これらを対象分析の手法として用いることには、まったく違和感はない。実際、これらは文句なしに有効な対象記述、分析のための手段であり、私は非生成言語学者がこの手の解析手法を毛嫌いする理由が不思議でならない。手法がすぐれているかどうかは、どの学派がそれを使っているかには依存しない<sup>21)</sup>。能力が十分でなく、やろうと思ってもうまくやれないというなら、まったく話は別であるが、“移動”や“削除”や“付加”や“代入”を使わずに、(非)容認可能性の反省を行わずに(「写像がうんたら」とか言う疑似科学的な議論に終止するような)意味解析を実践することは、私には顕微鏡なしで微生物学を、望遠鏡なしで天文学をやるような愚行にしか見えない。

### 3.3 (語の)意味はどこから来るか

変形をヒューリスティクスとして利用することを嫌い、逸脱例を考慮に入れないで正例のみを分析し、有意味性の相対化を怠った意味の研究をすることは、有意性の空間全体を考慮しないということである。これは怠惰な言語学者にありがちな研究姿勢であるが、このやり方に固執することには、一つ、非常に大きな限界がある。それは、

<sup>21)</sup>ただし、すぐれた分析手法を使う学派が、常にすぐれた解釈を提供するわけではない。少なくとも、生成言語学に関しては、このことは成立しない。

ということである。

私が見る限り、非常に多くの言語学者が、どの意味がどの語彙的要素から発生するのかに関して、驚くほど無頓着である。この傾向は、認知言語学の研究者と言えども、まったく変わらない。

この無頓着から来る弊害で非常に顕著なのは、次の二つのことである:

- (12) a.  $S = y_1 \cdots y_k \cdot x \cdot y_{k+1} \cdots y_n$  を文とするとき、要素  $x$  が、周囲に共起している“環境要素”  $\{y_1, \dots, y_n\}$  との相互作用から“獲得”した意味が、 $x$  の“固有の意味”だとして記述される
- b. 非語彙的な意味を語彙的要素に“無理やり押しつける”

これらの傾向はあまりに顕著なので、この誤謬を冒していない研究を探す方が困難なくらいだ。

### 3.4 語彙的意味の相互作用の存在

(12a)の問題を一言で言うと、語彙的意味の相互作用の存在が認識し損なわれているということである。実際、(12a)の獲得された意味とは、Langackerの用語で言えば意味の(相互)調節 (semantic (mutual) accomodation) [28, 29]の結果である。

(12a)の実例で典型的なのは、例えば、ある研究者が「格助詞“-で”の意味分類、記述をした」と称しておきながら、その人が実際にやったことは、“ $X$ で”の意味分類、記述で、結局、それは(“-で”ではなく) $X$ の意味分類、意味記述にすぎない、という場合である。この悪癖は学派に抛らずあまりに横行しているので「あたり前」のように見えるが、まったく妥当ではない。

「先日の台風で家を壊した」の“-で”は原因格で原因性を現わすのか、あるいは「先日、インターネットのオークションで詐欺にあった」の“-で”のように、時期・機会を表わすのかとか、あるいはまた「長

年の無理から体を壊した」の“-から”が源泉格なのか、原因格なのかとか、こういう厄介な問題には事欠かない。だが、通常は、記述的に十分だと思われる用法を幾つも列挙して終わりである。

このような問題は、もちろん、認識されていないわけではない。山梨 [47, 48] では、このような格解釈の「揺らぎ」を積極的に認定し、認知格の概念の下に再規定試みているが、なぜそうなるのかの説明は与えられていない。

山梨は [CONCRETE], [MANIPULABLE] などの(連続値を取る)意味素性への分解によって認知格を表示しようとしている。これは革新的な方法であり、高く評価したいが、一つ一つの意味役割 = 認知格の認定にしか妥当ではない。私の複層意味フレーム分析 (Multilayered Semantic Frame Analysis: MSFA) [44] が正しいならば、格の揺らぎの問題で本質的なのは、一つ一つの格解釈の曖昧性ではなく、意味役割の複層性、多次元性である。

### 3.5 還元主義の弊害: 生兵法はケガのもと

だが、このような格(助詞)の用法が有限個しかないという保証はない。格文法 [12, 13, 14] の理論はそれを要請するが、それが事実にもあったものかどうかは別問題である。

この問題は、表層格は有限個の基本格から「合成」されると考えるとうまく解消できる。実際、表層格が幾つかの異なる意味格の“合成” — 正確には多層的な重ね合わせ — であるという定式化は私の提唱している複層意味フレーム分析 [44] の解釈としては、妥当なものである。この場合、基本格は深層格相当のものである理論的必然性はないし、そのような前提は不要である。だが、そのような試みは、格の問題が非常に長い間に渡って言語学の特権的な研究対象であったにも係わらず、複層意味フレーム分析以外には存在しないようである。この理由はおそらく、多くの言語学者は、事実を正確に記述するための記述モデルを構築するより、自分の好みの事実を「説明」する説明モデルを寵用するため、事実をちゃんと見ないで、すぐに有限個の、なるべく数の少ない要素へ事実を押しこめようとする悪しき還元主義に奔るからだろう。多くの言語理論にとって、説明モデルに入り切らない現象は、「公式」には存在しない

のである。

### 3.6 意味の相互作用: これまでの言語学的意味論で、ほぼ完全に見落とされているもの

このことの是非はともかく、格助詞の解釈の揺らぎの問題に戻ると、これらは例外なく、

- (13) a. 第一に“-で”, “-から”の意味を“ $X_1$ で”, “ $X_2$ から”の意味から分離していないことから,  
b. 第二に“ $X_1$ で”, “ $X_2$ から”の意味を“ $X_2$ ”, “ $X_2$ ”の意味から分離していないことから

生じる“疑似問題”である。

実際、(深層)格、意味格、認知格の問題は、格助詞の問題ではまったくなく、助詞の付属してる名詞  $X_1, X_2$  の、文中のほかの要素との意味の相互作用 — Langacker 流の言い方をすれば意味の相互調節 [28, 29] — の問題である。

すでに言及したように、この問題には、私が開発した複層意味フレーム分析 [44] が単純明快で自然な解決を与えているはずである。

### 3.7 非語彙的意味の存在

(12a)の問題を一言で言うと、非語彙的意味の存在が認識し損なわれているということである。

(12b)の実例で典型的なのは、例えば、ある研究者が「終助詞“-た”の意味分類、記述をした」と称しておきながら、その人が実際にやったことは、“ $X$ だった”の意味分類、記述で、結局、それは(“-た”ではなく)“ $X$ だ”の意味分類、意味記述にすぎない、という場合である。この悪癖も、学派に抛らずあまりに横行しているので「あたり前」のように見えるが、まったく妥当ではない。

これがおかしいのは、それが

- (14) 発話は常にある目的のために行われるが、その目的は(少数の例外を除いて)語彙的にはコードされない

ということを反映していないからである。(14)に言い表されているのは、発話の発語媒介行為(性)である。

発話媒介行為 (性) の発見は Austin の発話行為論 [3] —並びに Goffman の社会学的観点からの言語分析 [17, 18, 19, 20]— によってもたらされた認識的革命的成果でも最大級に重要なものだと思うが、このことが言語学的記述に対して正確に何を意味しているか自覚している人は少ないように思う。

発話媒介性が意味しているのは、

- (15) 言語に発話媒介性があるということは、発話  $U$  は、それを音韻的に実現する文  $S$  を構成する語の語彙的意味には帰着できない意味をもつ

ということである。このことを自覚してないことから、先ほどの例であれば、一部の“-た”(e.g. “あ、(ここに)あった!”)には“発見”の意味があるとかいう(およそ荒唐無稽な)結論が導かれるのである。発見の“意味”があるのは“あ、(ここに)あった”のような文、正確には発言(の行動)であって、その発見の意味は、いかなる形態素にも還元しえないものである。

## 4 日本語の研究を部外者の目で評価する

以上のことから自ずと明らかなのは、次のことである:

- (16) 日本語、英語に限らず、意味の問題は難問であり、それをうまく扱うためには、データ駆動型の意味の実証的研究が必要だ。
- (17) 実証的な言語研究の基礎作りのために、統語に限らず、日本語のコーパスの利用価値を上げるような解析法の開発が必要だ。
- (18) だが、そのためには品詞ベース、解析ベースで間接的に意味を扱うのではなく、インデクス方式で意味の実体を直接的に扱う手法の確立が必要だ。

原理的には品詞タグづけと同様のことが意味タグづけでできればよい。だが、それは困難な課題である。課題事態がうまく定義されていない。なぜなら何が正解なのか、不明確だからである。文の意味を表わすということが、今だに何をどうすることなのか、よくわかっていないのだ。だから、まずそれが

ら始めなくてはならないのは明らかだ。この点で、Berkeley FrameNet [22] の取り組みは非常に示唆的だった。

### 4.1 日本語の「使える」研究が進んでいないわけ

先ほども述べたことだが、私は日本語の研究が好きではなかった。なぜか? それは先行研究を読むのが苦痛だったからである。私は国語学、日本語学、新日本語学の研究書、研究論文に目を通すのが嫌いであつた。それに横溢する人文科学ぶりに、正直に言って辟易していた。臆測の羅列、思いつきの殴り書きが論文をよい論文と見なす伝統に、私は苛立った。私は経歴はそうではないが、大学入学のために「文転」した根っから理系の人間であり、構造のない文章を読むのが非常に苦痛なのである。

私は嚙齧を覚悟で断言する: 分野を問わず、従来の日本語の言語研究で、科学的な意味で一定の記述的レベルに達している研究は少ない。説明的レベルに関しては「何をか謂んや」である。

#### 4.1.1 五里霧中

国語学、日本語学の研究は、とにかく見通しが悪い。どこに連れて行かれるか知らされないままで、真っ暗闇の中をあちこち連れ廻される感じがする。経験したことのある人ならわかることだが、これは非常に不安で不快な状況である。

国語学、日本語学の研究に重要な観察、洞察がないというわけではない。そういうものは、おそらく非常に豊富なのだと思う。路傍の石に思いがけなく美を見出す喜び—そういうものを求めている人々にはよい憩いの機会となるだろう。

だが、研究は詩作ではない。牧歌主義を脱却できない研究は、科学的価値を形成しない。ある対象の自然史 (natural history) づくりは科学的な意味での記述ではない。

#### 4.1.2 記述の標準的化の必要性

どういうことかということ、人文系では有益な意見交換、情報交換のために不可欠な記述フォーマット



が確立されていないし、それ以前にその必要性が自覚されてすらいないうことである。記述の内容、手法の標準化が行われていないので、大半の研究は何の資源にもならず、ムダに行われるのである。

何が、どう記述されなければならないかを定める記述のフォーマットがない限り、部外者から見た利用価値は低い。外部利用の可能性が低いほど、反証可能性は下がり、結果的に科学的な価値も下がる。確証バイアスは堅固なので、研究コミュニティ内部での議論のみから誤った学説が正されることは、実際には多くない。異なる研究活動と接することで初めて正される機会の発現した誤りというのは、歴史的に見ても非常にたくさんある。

#### 4.1.3 見通しのない記述主義を越えて

記述主義は重要である。だが、それは十分ではない。記述には見通しがなければならない。見通しのない記述主義は、単なる瑣末主義の別名である。

しかし、明らかに従来の日本語研究には徹底的に見通しという側面が欠落していた。その原因の一つは研究手法の非科学性、もう一つは応用を考慮に入れることの欠如である。私はそのことが実感として理解できるし、これこそが、多くの日本語研究者が反動で、生成言語学(の得てして下らない分析)に奔った理由ではないかと思う。

実際、国語学、日本語学のいわゆる「伝統的な研究」は—英語学から派生した理論先行型の日本語の言語研究と異なり—事実考証的な、データ基盤の研究でよいのだが、これとは別の理由で「見通しのない瑣末主義」に陥りがちであり、科学的な意味で実証的な研究とは言いかねる面がある。

これが招くのは、言語学研究論文、研究書を読み解くには、しばしば非常に苦痛が伴うという事態である。単に内容が難解であるという以上に、用語法の正確な定義もなく、往々にして意味不明で混乱した言明に狭間に埋もれており、読めば読むほどわからなくなるという経験をすることが少なくない。見通しのない些末な議論、脱線した議論、事実の観察よりも権威に縋った議論、そういうもので溢れている。更に困ったことに、そういうどうでもよいことを饅めているのが「よい論文」だと勘違いされている傾向がある。

#### 4.1.4 よい研究論文の条件

これに対し、私が理系的な基準で考える「よい論文」とは、以下のような特徴をもつものである：

##### (19) 研究論文は

- a. 論旨を含めて、内容が明快、明確であればある程よい
- b. 短かければ短いほどよい
- c. 記述が簡潔であればある程よい
  - i. 独自の用語は使わなければ使わない程よい
  - ii. 独自の用語、あるいは普通の用語を独特な意味で使っている場合、それに明快な定義が与えられていればいる程よい
- d. 何をどう分析したら、どんな結果が得られるのかが明示されていればいる程よい
- e. 研究により何が明らかになり、何が明らかになっていないかを明示するものであればある程よい
- f. これらの評価は客観的であればある程よい

このような基準から外れれば外れるほど、部外者にとっては利用価値が下がる。この点が文系の研究者にどれほど理解されているのか、私は根本的に怪しく思う。大勢の研究者が理論と応用の区別の上にあぐらをかいている。

## 4.2 無理に洞察を求めず、s/n 比を問題にしよう

この基準で見ると、明らかに従来の日本語研究の多くは、部外者から見た利用価値がゼロに近い。だから、日本語研究の大半は—それらにどんなに偉大な洞察が含まれていても—科学的には無価値同然なのである—少なくとも私にはそう思える。

ただ、これは言語学特有の問題というより、いわゆる人文系に全般的に認められる問題である。人文系の研究者は、研究というものが何であるか、それ自体が何であるかが解っていないように思う<sup>22)</sup>。

<sup>22)</sup>例えば、文系人間の書く文章の多くには「構造」がない—例えば問題の提起、関連する事実の記述、対応策の提案とその評価という一定の「議論の型」がない。そういうものがない状態で、現象の万華鏡を見せられる。これは事実そのものを偏愛している人でない限り、堪えられない。

何が問題かを一言と言うと、信頼に値する記述と信頼に値しない記述の比、妥当な記述と妥当でない記述の比、一種の  $s/n$  比が問題なのである。研究資料の  $s/n$  比が小さいうちは、誤りは「玉に瑕」程度だが、その比が一定値を超えると、それは価値のある研究資料とは見なしえない。

#### 4.2.1 ワザ比への伝統から脱却せよ

この辺に関する私の見解は、[43, 41, 42] に書いたことがあるので、ここでは繰り返さない。要点のみを言うと、研究はワザ比ではなく、共益だということである。

だが、非常に困ったことに、人文系の研究者は研究をワザ比へと勘違いする。その結果として、従来の言語学者たちによる日本語の研究は  $s/n$  比が低く—研究内容はともかく—研究資料としての価値が低くなっているのは、明らかである。

近代ヨーロッパで数学がいかに進歩したか、あるいは、当時は世界の最先端だった関孝和の「和算」が停滞した背景を見ると、ワザ比への弊害に関して面白い示唆が得られよう。「タルターリアへのカルダーノの裏切り<sup>23)</sup>」によってヨーロッパの数学はワザ比への伝統から脱却した。これによってヨーロッパの数学は一気に進歩し、近代数学が成立した。日本の数学者は最後までワザ比への伝統から抜け切れなかったことで、当時の世界最先端だった和算は閉塞し、自ずから過去の遺物と化した。

#### 4.2.2 他人の手間を考えよう

人文系の研究者はそうではないかも知れないが、研究者というのは基本的に仕事に追われており、使えるか使えないか解らない内容を暗号解読している暇はない。そんなムダな徒労に終わる可能性のある危険な努力をするより、自力でゼロから土台を作り直した方が早い。

実際、これは多くの自然言語処理系の研究者が採っている道である。彼らが文献で言語学関係の研究書をあげているのは、「いちおう、こういうのにも目は通しました」という以上の意味はない。夢々、彼ら

<sup>23)</sup>これが裏切りだったのか、そうでなかったのには諸説あるようだ。私は [37] や [36] を読んで、裏切り説はウソではないかという考えに傾いた。

が言語学の研究を「参考にして」研究をしているとは思わない方がいい。

## 5 これからの言語学: 言語学から言語科学へ

言語学は単に定義によって言語の科学であるばかりでなく、実質的に言語科学にならなければならない。

### 5.1 言語処理の基礎科学としての言語学?

私が想像する限り、実質的な言語科学が自然言語処理に対する関係は、理学部の物理学科、化学科で行われている研究と工学部で行われている研究の関係、理学部の生物学科で行われている研究と農学部で行われている研究の關係に相当する。この対応を想定する限り、言語学科が文学部に存在し、主に人文系の研究者によって担われているという現状は、非常に奇妙なことだ。

これから一つ非常に重要なことが帰結する: 世の中の言語学者連中の多くは、私の言う言語科学者ではない。彼らはせいぜい、言語科学者の「なり損ない」だ。

実際、この乖離が今の言語学を自然言語処理の研究者から見て「使えない」ものになっている理由の一つである。私は反論を承知で断言するが、言語学が人文系の学問である時代は終りつつあると考え、構造改革を始める必要がある。そうでないなら、言語学は本当に無用な学問になり果てるであろう。

自然言語処理 (NLP) が行き詰まっている理由—特に最後の難関である意味処理の分野で行き詰まっている理由—の一つは、ここにある。NLP は、その性質から言って実用志向の強い、応用分野である。だが、それに十分に信頼性のある基礎研究を提供する分野が存在しないのである。それは本来ならば言語学に期待されてしかるべきことなのだが、その期待が報われないのは、もう日本の NLP の業界では完全に知れ渡っている<sup>24)</sup>。

<sup>24)</sup>欧米の場合は、まったく話が違う。計算言語学 (Computational Linguistics) という分野が—NLP と半ば相乗りの状態だが—確立しており、基礎と応用の分離は、日本以上にずっと進んでいるように見受けられる。GPSG [15], TAG [2] に始まり、HPSG [31], LFG [4], Generative Lexicon Theory [32, 33] がその流れに乗っている。ただ一つ、私には不思議でならないのは、JPSG [21] がなぜ日本で定着しなかったのかである。HPSG に合流しただけなんだろうか?

## 5.2 何のための言語研究?

話は大幅に本論からズレるが、私は人文系の研究者、教育者にありがちな「自分たちのやっていることは役に立たなくてイイ」という態度が大嫌いだ。そんな向上心のないこと言うくらいなら、いっそのこと講義をするな。学生を取るな。税金を使うな。別に実生活に今すぐに役立つ必要などないけれど、未来永劫に渡って誰かの役に立つ可能性のないような研究は、まったく無意味だ。共益の見こみのない研究など、存在価値はない。

だが、(創始者の)理念から見て応用を軽視する傾向が避けられない生成言語学ばかりでなく、その対抗馬であるはずの認知言語学であっても、その分野で「研究」に従事している人々に本当に役に立つ研究をしようという心がけをもった研究者は皆無に等しい。彼らの最大の関心事は、言語学という狭い世界の内部での派閥争いで勝者になることである。私が見る限り、自分の研究分野の実質的進展は、二の次である。

このような事情もあってか、非常に残念なことに、いわゆる「言語学者」の多く—特にその卵たち—は就職のためだけに「研究」する<sup>25)</sup>。比較的活発に研究発表をするのは、就職前の数年だけで、就職するとぱったり研究も発表も辞めてしまう人が多いように思われる。これを説明するのに、しばしば「大学での業務が忙しく、家事との両立が大変で研究する時間がない」という理由を耳にする。私は、これはまったく言いわけになっていないと考える。そういう人たちは家で何をしているだろうか? たいいていの人は研究をする時間はないが、家でテレビを見る暇がある。その時間に研究できないのだろうか? 人には誰でも休息が必要だと言うかも知れない。仕事をバリバリこなす人がそんな苦言を漏らすのは聞いたことがない。

これは要するに、覚悟の問題、正確に言えば、研究者としての人生設計の問題なのだと私は考える。それがうまくできる人は研究できるし、うまくできない人は研究を続けられない—本当はそれだけの話である。研究者は、当然のことだが、見通しを立て

<sup>25)</sup>ただし、これは言語学に限ったことではない。言語学に関して以上なのは、就職先が語学教師であることだ。これは言語学では研究の質に十分な選択圧がかかっていないことを意味する。就職の有利・不利を決めるのは、得手して研究の質ではなく、単なる学派色だからだ。

てなるものである。例えば私は、将来的に研究の邪魔になる可能性のあることから、可能な限り身から遠ざかる努力をして来た。今でも私は研究の時間を作るためにテレビは見ない(私はそもそも、テレビセットをもっていない)。それで何も困らない—少なくとも研究者としては、巷で流行していることを知らなくて困ることは、滅多にない<sup>26)</sup>。それは私にはまったく自然なことだった。私は研究者にはリスク覚悟でなるものだというのを高校生の時分から知っていた。

こういうことを言うと、しばしば「平均的な人間はそうじゃない」と「反論」される。だが、私に言われれば、そんなことは、全然まったく問題じゃない。このように批判する人々は研究が何であるか、根本的に誤解していると思う。何で研究者の基準を、いわゆる「平均」に合わせる必要があるのだろうか? すぐれた研究者がすぐれた成果を出すために、他人の见えないところでどれほどの努力をしていると思っているのだろうか?

研究者というものに関して、もっと強調されるべきなのに世間的には強調されないでいない—というより自分の分野を実際以上に魅力的に見せかけるために隠蔽されている—ことがある。それは研究者になるという選択は、相当にリスクの高い選択であるということだ。研究者というのは、リスクを負う覚悟のある者だけがなるべきものだと私は思う<sup>27)</sup>。

<sup>26)</sup>学生のウケを取るのに必要だと言う人がいる。そういうことを言う人は、自分の研究の内容を学生に対して魅力ある語り口が説明できない人が多いように思う。流行を共有することを人気取りに使うことは、諂いである。そんなことをするより、学生に対して知的に挑戦した方がポテンシャルのある学生には喜ばれど私は思うし、学生に対して知的に挑戦しない教官は、教官として失格だと私は考える。講義に出席している学生の全員にウケる必要など、はじめからないと私は考える。

こうこうことを言うとしばしば「エリート主義」だと言われるが、それは完全に筋違いである。教官が生徒に阿るようになれば、教育は墮落する。

<sup>27)</sup>因みに、私はどれぐらいの覚悟があるかと言うと、例えば私は普通の家庭をもつことずいぶん前から放棄している。これには一つ、重大な理由がある。私は浪人生の時に読んだ大江健三郎の著作から、家庭をもっていると自分の主義主張を貫くことが難しくなる場合があることを知った(大江氏は自分の書いた小説の内容を巡って右翼から受けた。家庭を守るために言いたいことが言えない状況に追い込まれた彼の苦痛は痛いほどよくわかった)。ここまで極端ではないにせよ、家庭をもつことは、世間的な幸福と引換えに「責任」という不自由を背負いこむことである。それをした人はいしてよいと思うが、私は嫌だ。人生において、何を選択することは、別の何かを放棄することである。私は最大限に思考と表現の自由を選択したいので、それと両立しないものは選択しないようにしている。

こう言うと、モノのわかってない人々から「お前は無責任だ」という批判を受ける。断っておくが、ここで「無責任」という概念はもちだすのは、まったくの筋違いである。私は責任を負う必要が生じないように努力しているのであって、責任(や義務)を回

## 6 おわりに

どんな理由があるにせよ、日本語の統語論を「使える」ようにするには、土台からやり直されなければならないのは明らかだ。統語論をやり直すだけでなく、マトモな意味論もやり始めなければダメだ。イメージ(スキーマ)だとかメタファーだとかが言語の意味の記述だと思っているようでは「使える」意味論はいつまでも始まらない。形式意味論(Formal Semantics)の記述可能性の「壁」を乗り越えなければならないのは当然だが、まったく新しいことなど、誰にも始められない。だから、形式意味論の限界を知りつつ、それとの互換性を画策するのは重要なことである。認知言語学の(無責任に楽天的な)指導者連中のように、形式意味論の一切切を切り捨てて平然としても、意味論全体の実質的な進歩には繋がらない<sup>28)</sup>。形式意味論に限界があるのは、誰でも—形式意味論をやっている人たち自身でも—知っている。それを見捨てて無関係なことを始めるのは、やろうと思えば、誰にでもできる。難しいのは、過去の遺産を放棄せず、それを取り込むような、新しい記述の可能性を切り開くことである。

私は自分がそのような新しい可能性を提供する能力があるとは思わないが、その実現のために何が必要であるかは知っているし、そのために何かができると感じた。それが日本語の意味(役割)タグつきコーパスの開発である。意味タグづけの体系、並びに意味タグつきコーパスの開発は、それ自体では新しい成果を生み出すものではないが、言語学を言語科学に質的に変換するという、非常に困難な仕事を始めるためのインフラ構造なのである。それが統語構造理論のインフラにもなるのは、統語構造がどんな情報を反映し、どんな情報を反映しなくていいかを明示するからである。その区別なしで日本語の統語論をやろうとすれば、うまくいかないのは目に見えている。英語の統語論と同じような“透明”な統語論を構築するやり方は、日本語のような言語には通用し

避しているのではない。私は自分が(日本という国家のメンバーとして)家庭をもつ義務があるとはまったく考えない(人にはせいぜい結婚する権利があるだけだ)。私は強くリバイアサニズムを信じているので、自分の人生で自分が何をやるかは、(法と秩序が許す限りで)自分の勝手に選べると思っている。私が存在するのは、私自身のためである。だが、こういう人生観をもつ者は日本という国ではまったく得をしない。理由は[34, 35]などにある通りだ。

<sup>28)</sup>このような切り捨て主義の傾向は、[27]のような書物に顕著である。

ない。それはもう、過去数十年間のチョムスキー派生成言語学の失敗から明らかなのだ。

## 参考文献

- [1] A. Abeillé, editor. *TreeBanks: Building and Using Parsed Corpora*. Kluwer Academic Press, 2003.
- [2] A. Abeillé and O. Rambow, editors. *Tree Adjoining Grammars: Formalisms, Linguistic Analyses, and Processing*. CSLI Publications, Stanford, CA, 2000.
- [3] J. L. Austin. *How to Do Things with Words: The William James Lectures Delivered at Harvard University in 1955*. Harvard University Press, Cambridge, MA, 2nd edition, 1962. [邦訳: 『言語行為論』(坂本百大 訳). 勁草書房].
- [4] J. Bresnan. *Lexical-Functional Syntax*. Blackwell, 2001.
- [5] J. L. Elman. Grammatical structure and distributed representations. In S. Davis, editor, *Connectionism: Theory and Practice*, Vancouver Series in Cognitive Science, Vol. 3, pages 138–178. Oxford University Press, Oxford, UK, 1992. [Reprinted in Geirsson, H. and Losonsky, M. (Eds.), *Readings in Language and Mind*, 526–558. Basil Blackwell].
- [6] G. R. Fauconnier. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, MA: MIT Press, 1985.
- [7] G. R. Fauconnier. *Mappings in Thought and Language*. Cambridge, MA: Cambridge University Press, 1997.
- [8] G. R. Fauconnier and M. Turner. Conceptual projections and middle spaces. Cognitive Science Technical Report (TR-9401), Cognitive Science Department, UCSD, 1994.
- [9] G. R. Fauconnier and M. Turner. Blending as a central process of grammar. In A. D. Goldberg, editor, *Conceptual Structure, Discourse, and Language*. CSLI Publications, 1996.
- [10] G. R. Fauconnier and M. Turner. Conceptual integration networks. *Cognitive Science*, 22:133–187, 1998.
- [11] C. Fellbaum, editor. *WordNet: An Electronic Lexical Database*. MIT Press, 1998.
- [12] C. J. Fillmore. The case for case. In W. Bach and R. T. Harms, editors, *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, 1968. [Reprinted in Fillmore (2003), *Form and Meaning in Language, Vol. 1: Papers on Semantic Roles*, pp. 23–122. CSLI Publications.].
- [13] C. J. Fillmore. Some problems for case grammar. In R. O'Brien, editor, *Report of the Twenty-Second Round Table Meeting on Linguistics and Language Studies*, pages 35–56. Georgetown University Press, Washington, D.C., 1971. [Reprinted in R. Dirven and G. Rodden, eds. (1987). *Fillmore's Case Grammar — A Reader*, pp. 59–69. Heidelberg: Julius Groos Verlag].
- [14] C. J. Fillmore. The case for case reopened. In P. Cole and J. M. Sadock, editors, *Syntax and Semantics, Vol. 8: Grammatical Relations*, pages 59–82. Academic Press,

- New York, 1977. [Reprinted in Fillmore (2003), *Form and Meaning in Language, Vol. 1: Papers on Semantic Roles*, pp. 175–199. CSLI Publications.].
- [15] G. Gazdar, E. Klein, G. K. Pullum, and I. A. Sag. *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford, 1985.
- [16] D. Gildea and M. Palmer. The necessity of parsing for predicate argument recognition. In *Proceedings of ACL 2002*, 2002.
- [17] E. Goffman. *Frame Analysis*. New York: Harper, 1974.
- [18] E. Goffman. Response cries. *Language*, 54:787–815, 1978. [Reprinted in Goffman 1981, pp. 78–123].
- [19] E. Goffman. Footing. *Semiotica*, 25:1–29, 1979. [Reprinted in Goffman 1981, pp. 124–159].
- [20] E. Goffman. *Forms of Talk*. University of Pennsylvania Press, Philadelphia., 1981.
- [21] T. Gunji. *Japanese Phrase Structure Grammar: A Unification-Based Approach*. Reidel, 1987.
- [22] C. R. Johnson and C. J. Fillmore. The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In *Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics (ANLP-NAACL 2000)*, pages 56–62, 2000.
- [23] P. Kingsbury and M. Palmer. From TreeBank to Prop-Bank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*, 2002.
- [24] P. Kingsbury, M. Palmer, and M. Marcus. Adding semantic annotation to the Penn TreeBank. In *Proceedings of the Human Language Technology Conference*, 2002.
- [25] K. Kuroda. *Foundations of PATTERN MATCHING ANALYSIS: A New Method Proposed for the Cognitively Realistic Description of Natural Language Syntax*. PhD thesis, Kyoto University, Japan, 2000.
- [26] G. Lakoff. *Women, Fire, and Dangerous Things*. University of Chicago Press, 1987. [邦訳: 『認知意味論』(池上嘉彦・河上誓作訳). 紀伊国屋書店.].
- [27] G. Lakoff and M. Johnson. *The Philosophy in the Flesh*. Basic Books, 1999.
- [28] R. W. Langacker. *Foundations of Cognitive Grammar, Vol. 1: Theoretical Prerequisites*. Stanford University Press, 1987.
- [29] R. W. Langacker. *Foundations of Cognitive Grammar, Vol. 2: Descriptive Applications*. Stanford University Press, 1991.
- [30] M. P. Marcus, B. Santorini, and M. A. Marcinkiewicz. Building a large annotated corpus of english: the Penn Treebank. *Computational Linguistics*, 19(2):313–330, 1993.
- [31] C. J. Pollard and I. A. Sag. *Head-driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Center for the Study of Language and Information/The University of Chicago Press, Stanford, CA/Chicago, IL, 1994.
- [32] J. Pustejovsky. The generative lexicon. *Computational Linguistics*, 17(4):409–440, 1991.
- [33] J. Pustejovsky. *The Generative Lexicon*. MIT Press, 1995.
- [34] K. van Wolferen. 日本/権力構造の謎(上, 下). 早川書房, 1994. Original title: *The Enigma of Japanese Power: People and Politics of a Stateless Nation*. Vintage, 1990.
- [35] K. van Wolferen. 人間を幸福にしない日本というシステム. 新潮社, 2000. Original title: *The False Realities of a Politicized Society*. ??, 1994.
- [36] L. ムロディナウ. たまたま: 日常に潜む「偶然」を科学する. ダイヤモンド社, 2009. 原典: Leonard Mlodinaw: *The Drunkard's Walk: How Randomness Rules our Livers*, Vintage, 2008.
- [37] M. リヴィオ. なぜこの方程式は解けないのか? 天才数学者が見出した「シムメトリー」の秘密. 早川書房, 2007. 原典: Mario Livio: *The Equation that Couldn't be Solved: How Mathematical Genius Discovered the Language of Symmetry*, Simon & Schuster, 2005.
- [38] 徳永 健伸. 辞書と情報処理. In 単語と辞書(言語の科学 3), pages 156–190. 岩波書店, 2004.
- [39] 池原 悟, 徳久 雅人, 村上 仁一, 佐良木 昌, 池田 尚志, and 宮崎 正弘. 非線形な重文複文の表現に対する文型パターン辞書の開発. 情報処理学会研究報告, NL-170(25):157–164, 2005.
- [40] 黒田 航. 偉大な女性 Elizabeth Bates の思い出. URL: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/memory-of-liz-bates.pdf>, 2003.
- [41] 黒田 航. 「そんなことは にすでに書いてある」というありがたいお説教について: 研究資料の s/n 比を測定しよう. URL: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/sn-ratio-of-linguistics.pdf>, 2004.
- [42] 黒田 航. 作例中心主義を脱却しよう: 言語科学における作例の意味について. [URL: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/on-making-up-examples.pdf>], 2004.
- [43] 黒田 航. 「説明強迫症」の治療と予防: 言語学の「緩やかな科学化」のための試論. [<http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/against-explanation-impulse.pdf>], 2005.
- [44] 黒田 航 and 井佐原 均. 意味フレームを用いた知識構造の言語への効果的な結びつけ. 信学技報, 104 (416):65–70, 2004. [増補改訂版: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/linking-1-to-k-v3.pdf>].
- [45] 黒田 航 and 井佐原 均. 日本語の意味タグ体系を定義する試み: FrameNet の視点から. In 言語処理学会第 10 回年次大会発表論文集, pages 148–151. 言語処理学会, 2004. [増補改訂版: <http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/jfn-nlp10-rev.pdf>].
- [46] 永田 昌明. 形態素解析. In 単語と辞書(言語の科学 3), pages 54–92. 岩波書店, 2004.
- [47] 山梨 正明. 格の複合スキーマモデル: 格解釈のゆらぎと認知のメカニズム. In 仁田 義雄, editor, 格をめぐる, pages 39–65. くろしお出版, 1993.

- [48] 山梨 正明. 日常言語の認知格モデル (1)–(12). 言語, 23(1月号–12月号), 1994.
- [49] 影山 太郎. 文法と形態論. In 単語と辞書 (言語の科学 3), pages 2–51. 岩波書店, 2004.
- [50] 黒橋 禎夫 and 長尾 眞. 並列構造の検出に基づく長い日本語文の構文解析. 言語処理, 1(1):35–58, 1994.
- [51] 増市 博 and 大熊 智子. Lexical Functional Grammar に基づく実用的な日本語解析システムの構築. 自然言語処理, 10(2):79–109, 2003.
- [52] 佐藤 理史. 境界認定の提案 (1) コンセプトと実現法. 情報処理学会研究報告, 2004-NL-164:25–32, 2004.
- [53] 佐藤 理史. 境界認定の提案 (2) 背景と思想. 情報処理学会研究報告, 2004-NL-164:33–40, 2004.