

知能からみた日本語と英語の普遍文法

敬和学園大学

安藤司文

概要

人間は、質問と応答、知識の獲得、物語の理解などの知的な活動を、頭の中にある概念構造を用いて行っている。人間の頭の中にある概念構造を記号化したものを普遍言語と名付け、その普遍言語をベースにして、日本語と英語の普遍的な言語構造と個別言語固有の言語構造を検討し、次の点を明らかにした。

- (1) 日本語や英語などの自然言語の文はどれも“～がある”、“～である”、“～をする”という3種類の基本単文(基本意味単位)で構成されている。
- (2) どのような複雑な意味を表現する文でもいくつかの単文が結合されたものである。
- (3) 日本語や英語の単文の語順は、基本単文内での語順によって決定される。
- (4) 多様に表現される文でも、意味が同じであれば、同じ意味構造を持っている。

AN INTELLECTUAL LOOK AT UNIVERSAL JAPANESE AND ENGLISH GRAMMER

Shimon Ando

Keiwa College

ABSTRACT

Human beings seem to evolve intellectual activities, such as questioning and answering, acquisition of knowledge, comprehension of stories, etc, utilizing conceptual structures in their human brains. If these intra-cerebral conceptual structures can be extracted as symbolized in some particular concrete ways, we designate them UNIVERSAL LANGUAGE. We have studied universal linguistic structures and individual linguistic structures as they are inherent in Japanese and English, based upon the UNIVERSAL LANGUAGE. This has brought about the results as follows:

- (1) Natural sentences in Japanese and English are composed of three kinds of primitive sentences (basic meaning units), that is, {~ ga aru}, {~ do aru}, {~ wo suru}
- (2) Sentences, which even express complicated meanings, are unifications of simple sentences.
- (3) The word order in simple sentences in Japanese and English has been determined by that of the primitive sentences.
- (4) Sentences, expressed in various ways with the same meaning inherent in them, have the same conceptual structures.

1. はじめに

日本語から英語、中国語、マレー語などの色々な言語に、あるいは英語から日本語やその他の多くの言語に多元的に機械翻訳を行うためには、これらの言語に共通する普遍文法を解明する必要がある。文法理論はもともと言語を普遍的にとらえようとする学問であって、これまでいろいろな文法理論が提案されている(例えば、1, 2, 3)。しかし、これらの文法理論はまだ普遍文法として未完で、機械翻訳などで実用化できる状態ではない。

人間は自然言語を用いて、質問応答を行ったり、知識を獲得したり、推論を行ったりしている。人間の知的な活動を支えているメカニズムが自然言語の中に存在しているように思える。しかし、自然言語は、話し手の頭の中にある考えと同じ考えを聞き手の頭の中に想起させるための道具にすぎない。人間の知的な活動の本質的な役割を担っているのは、頭のなかにある概念構造であると考えられる。人間の頭の中にある概念構造を記号化して取り出したものを普遍言語と名付けて、これをベースに普遍文法の研究を行っている。図1に自然言語と普遍言語の関係を示すが、この役割を明確に認識することが普遍文法理解のポイントになる。



図1 自然言語と普遍言語との関係

人間が、“人間は死ぬ”という知識を獲得するプロセス、“太郎が花子にあげたよ、何を？バラです。何時何処で？今日新潟で”という質問応答から、“太郎が今日新潟で花子にバラをあげた”という内容を理解するプロセスは個別言語にとらわれない普遍的な知的なプロセスである。このような知的なプロセスを追求する過程で、本普遍文法(1~20)が生まれた。

個別言語やその文法から普遍文法の全体像を明らかにすることは難しいが、普遍文法から個別言語やその文法を見ると、どの部分が普遍的でどの部分が個別言語特有の言語や文法であるかが容易にわかる。ここでは普遍言語や普遍文法から、英語と日本語の言語構造を説明する。本報はこれまでの一連の普遍文法に関する研究(5~20)の総括である。また、これらの普遍文法を分かりやすく説明した本“映画で学ぶ知能からみた日本語と英語の普遍文法(スクリーンプレイ出版社)”(4)を出版したので、詳しくはその本を参照されたい。

2. 普遍文法は文構造文法

従来の文法と決定的に異なるのは、全ての文は文と文の組み合わせでできていると考えている点である。そのために、これを文構造文法(Sentence Structure Grammar S S文法)と名付けた。文は文と文が多層に縦結合(格結合)や横結合(論理結合)されて構成されているので、多層格論理構造文法とも呼んでいる。例えば、次の文

想像力がなかったならここで暮らしていくことはとてもできませんわ

I simply couldn't live here if I hadn't any imagination.

は英語も日本語も次の三つの文から構成されていると考える。

文1：私がここで暮らしていく。

I live here.

文2：私は(そうすること)はとてもできません。

I simply couldn't (do that).

文3：私に想像力がなかった。

I hadn't any imagination.

それが、次のように縦横に結合されたものである。

(私に想像力がなかった)なら→(私は(私がここで暮らしていくこと)はとてもできませんわ。)

(I simply couldn't (↓ I live here))←if(I hadn't any imagination.)

印で示された部分を表現禁止にし、上の文を下に埋め込み、"()"や"←, →, ↓"などの矢印を取り除き、助詞を若干変更させたものが自然言語である。日本語も英語も文1, 文2, 文3で構成されているが、文の順序と文内の語順と表現禁止になる単語が日本と英語では異なっている。文で構成されている全体の骨格は普遍的な部分である。日本語と英語の単文内での語順の違いについて後で説明する。

3. 人間はどのようにして知識を獲得するか

人間は知識を獲得する場合、まず仮説をたて、その仮説が実社会で有効であることを確認してはじめて、それが知識として認められることになる。知識獲得のためには、まず仮説を生成しなければならないが、例えば、"ハモンドさんが死んだ"という文から簡単に次のようにして"人間は死ぬ"という仮説が生成できる。"ハモンドさんは人間である"ことが分かっているならば、分かっていることは、自然言語ではいちいち表現しない(4)ので、"ハモンドさんが死んだ"は"人間であるハモンドさんが死んだ"という意味である。これは、次の文1と文2

文1: ハモンドさんは人間である。

文2: ハモンドさんが死んだ。

が、次のように縦に結合されたものである。

ハモンドさんは人間である
↓
ハモンドさんが死んだ。

ここで、文1と文2を論理助詞の"ならば"で横に結合し、"ハモンドさん"を変数のXに置き換え、動詞を現在形にし、格助詞を若干変更させると次のようになる。さらに"X"を表現禁止にすると下の右ようになる。

(Xが人間である)ならば→(Xは死ぬ) (人間である)ならば→(死ぬ)

つまり、<人間であるならば死ぬ>、<人間は死ぬ>となる。これは仮説であるが、現実の世界でありとあらゆる場合に成立するので、人間はこれを知識と認めている。このように普通言語から仮説は簡単に生成できる。

3. 文の意味構造

3.1 基本単文

一つの動詞で表現される文を単文と呼ぶ。どんな複雑な文でもいくつかの単文が縦横に結合されてきたものであることはすでに述べたが、一つの文であると考えられている単文も実は次に述べるいくつかの基本単文で構成されている。これ以上分解できない基本単文は意味の最小単位を表現するので、これを基本意味単位と呼ぶ。基本意味単位、即ち、基本単文は次の五つの格と三つの補助的な格で構成されている。

A格; 主格	Agent case	(誰) が	(who)	X格; 補助格	Auxiliary case
T格; 時間格	Time case	(何時) 何	(when)	Y格; 応答格	Yes_No case
S格; 空間格	Space case	(何処) で	(where)	Z格; 全体格	Zentai case
O格; 対象格	Object case	(何) を	(what)		
P格; 述語格	Predicate case	(どう) する	(how)		

X格は複雑な語順の文を生成するときや、英語では"it~that"などの構文で"that"以下を示すときなどに用いられる補助的な格である。Y格は"yes", "no", "はい", "いいえ"などの応答の時の単語を格納するのに用いられる。Z格は文全体の内容表現するとき用いる格である。A T S O Pが必要最小限の意味を表現するのに必要な格である。A格に"(誰) が (who)", T格に"(何時) (when)", S格に"(何処) で (where)", O格に"(何) を (what)", P格に"(どう) する (how)"を埋め込み、これと横関係(論理関係)で結合される"(何故) (why)"も含めて考えると、次のように、日本語も英語も基本意味単位は<I H 5 W>である。格助詞を中抜き文字で示す。

(何故) → [((誰) が (何時) (何処) で (何) を (どう) する)] (why) → [((who) (how) (what) (where) (when))]

五つの格の順番が日本語と英語と中国語では異なる。

<日本語の場合>

A+T+S+O+P (アトソプと読む)

<英語の場合>

A+P+O+S+T (アポストと読む)

<中国語の場合>

A+T+S+P+O (アトスポと読む)

P格に"(あ)る"を埋め込めば、"〜がある"という基本単文になる。P格に"(であ)る"を埋め込めば、"〜である"という基本単文になり、P格に"(す)る"を埋め込めば、"〜をする"という基本単文になる。つまり、基本単文には次の三種類が存在する。

" ~がある " PS-EX
" ~である " PS-IS
" ~をする " PS-DO

この三つの基本単文をPS-EX, PS-IS, PS-DOの記号で表す。PSは英語の(PRIMITIVE SENTENCE), "EX"は英語の(EXIST)の略で, "IS"は(IS), "DO"は(DO)である。これ以外の基本単文はない。ありとあらゆる文はこの三種類の基本単文で構築されている。" ~がある", " ~である", " ~をする"は次のような意味構造をもっている。各格に単語を埋め込むと、次のようになる。

$[(\text{何か})_A \text{が} (\text{何時か})_T (\text{どこか})_S \text{に} - (\text{あ} \text{る})_P]$	$[(\text{花})_A \text{が} (\text{今})_T (\text{水})_S \text{に} - (\text{あ} \text{る})_P]$
$[(\text{何か})_A \text{が} (\text{何時か})_T (\text{どこか})_S \text{で} (\text{何か})_O (\text{であ} \text{る})_P]$	$[(\text{花子})_A \text{が} (\text{今})_T (\text{ここ})_S \text{で} (\text{美人})_O (\text{であ} \text{る})_P]$
$[(\text{何か})_A \text{が} (\text{何時か})_T (\text{どこか})_S \text{で} (\text{何か})_O \text{を} (\text{す} \text{る})_P]$	$[(\text{太郎})_A \text{が} (\text{今})_T (\text{ここ})_S \text{で} (\text{それ})_O \text{を} (\text{す} \text{る})_P]$

これは"花が今ここにある", "花子は今ここで美人である", "太郎が今ここでそれをする"という自然文を表現したことになる。

3.2 準基本単文

ところで、<始点にあったものが、中間点を通って、終点にある>という状態にすることは、"移動"であるが、このような意味構造が基本的な意味単位として存在している。これを次のように表記する。

(始点) から → (中間点) → (終点) へ	
[(何か) が (何時か) () - (存在) する]	PS-SSS

<存在しなかったものが、存在する>という状態は"発生"で、これを次のような意味構造として表現する。

非(存在) → (存在)	
[(何か) が (何時か) (何処か) に - ()]	PS-E'E

<存在していたものが存在しなくなる>という状態は"消滅"で、これを次のような意味構造として表現する。

(存在) → 非(存在)	
[(何か) が (何時か) (何処か) に - ()]	PS-EE'

これは基本意味単位PS-EXを横に2~3個結合したものを略記したもので、これを準基本意味単位、あるいは準基本単文とする。これらをPS-SSS, PS-E'E, PS-EE'の記号で表す。単文の意味はこれらの基本単文と準基本単文によって、構成されている。これを"浮かべる"と"与える"を例にして説明する。

3.3 単文の意味構造

(1) "浮かべる"の意味構造

太郎が今ここで花を水に浮かべる。

は、次に示すように、PS-EXの基本単文1とPS-DOの基本単文2が縦に結合されたものである。

基本単文1: (花が今水にある)	PS-EX
基本単文2: (太郎が今ここで(そのよう)にする)	PS-DO

これを絵で示せば図2のようになる。この絵をそのまま記号や文字列に置き換えたのが図3で、これを普通言語としている。これが"浮かべる"の意味構造(意味フレーム)である。単語を埋め込む格をスロットと呼ぶが、単語がまだ埋め込まれていないスロットを、意味フレームで指示された順序で並べると次のようになる。

[(主語) (時間) (空間) で (目的語) (場所) に (浮かべ) る]

"浮かべる"という動詞は"(主語), (時間), (空間), (目的語) (場所)"などの変数を持っている関数である。このような穴(スロット)に"太郎", "今", "ここ", "花", "水"を埋め込んだのが、図3である。この意味構造(意味フレーム)すなわち、普通言語、で指示されている順番に単語と助詞を並べると、先に述べた文になる。これが日本語や英語などの自然言語と普通言語との関係である。

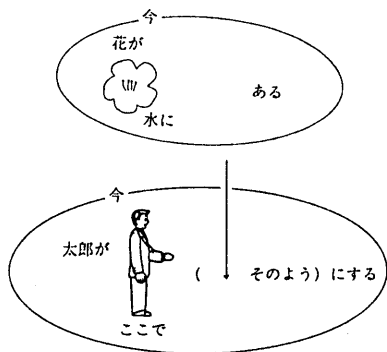


図2 “太郎が今ここで花を水に浮かべる”の意味構造

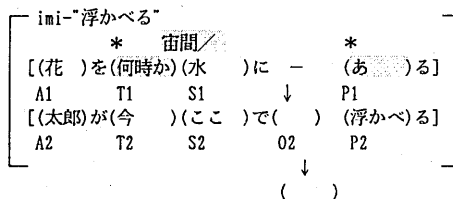


図3 “太郎が今ここで花を水に浮かべる”の意味構造

さらにもう一例，“与える”の意味構造（意味フレーム）について考える。

(2) “与える”の意味構造

移動を表す準基本単文PS-SSSをPS-ISのO2格に縦に結合し、それをさらにPS-DOのO3格に結合する。ここで(A3)で上から3番目の基本単文のA格のスロットを示すこととして、(始点)に(A3)と同じ単語を埋め込むことを(始点)=(A3)で表すことにする。(始点)=(A3)、(終点)=(A2)、(T1)=(T2)=(T3)、(S2)=(S3)とすると、“与える”の意味構造(意味フレーム)になる。ここで(A3)=(太郎)、(T3)=(今)、(S3)=(ここ)、(A2)=(花子)、(A1)=(バラ)とすると、図4のようになる。

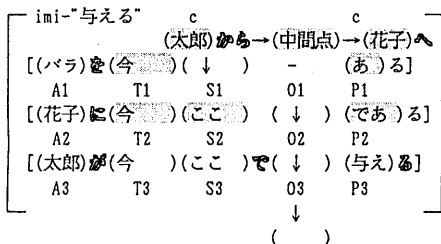


図4 “太郎が今ここで花子にバラを与える”の意味フレーム(概念構造)

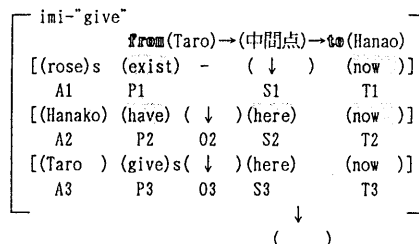


図5 英語の“Taro gives Hanako roses here now.”の意味フレーム

一番上のPS-SSSは<バラが今太郎のところから花子のところに移動する>ことを示し、二番目のPS-ISは<花子は今ここでそのような状態である、即ち、花子は今ここで花子のところにバラがあるという状態である>ことを示し、三番目のPS-DOは<太郎が今ここでそのような状態にする>ことを示している。PS-SSSの状態になった<時間>とPS-ISの状態になった<時間>とPS-DOの状態になった<時間>は同じ“今”である。PS-ISとPS-DOの<空間>は同じ<ここ>である。同じ単語はあちこちにある。英語の場合は図5に示すように、基本単文内での格の順番が異なるだけで、基本的には全く同じ意味構造である。この意味フレームで指示されている順番で表現可となっている単語と助詞を並べたのが、次に示す日本語と英語の自然言語の文である。

太郎が今ここで花子にバラを与える。

Taro gives Hanako roses here now.

意味構造は日本語も英語も全く同じである。基本単文内での語順によって単文の語順が決定される。これを次に説明する。

4. 基本語順

図4の意味フレームと図5の意味フレームの単語や記号を指示されている通りに、並べると次のようになる。

([(太郎)が(今)(ここ)で([(花子)に(今)(ここ)で([(バラ)を(今)((太郎)から→(中間点)→(花子)へ)-(あ)る]) (であ)る]) (与え)る])

([(Taro) (give)s (↓ [(Hanako) (have) (↓ [(rose)s (exist) -(↓ from(Taro)→(中間点)→to(Hanako)) (now)]) (here) (now)]) (here) (now)])

(), [, ↓ ”などの記号を取り除くと次のようになる。

太郎が今ここで花子に今ここでバラを今太郎から中間点花子へあるである与える
 Taro gives Hanako have roses exist from Taro 中間点 to Hanako now here now here now.

重複する単語は表現禁止にしなければならないので、それらを取り除くと次のようになる。

太郎が今ここで花子にバラを与える Taro gives Hanako roses here now.

日本語と英語の基本語順は次のようなる。

[日本語の場合]					[英語の場合]						
太郎が	今	ここで	花子に	バラを	与え	Taro	gives	Hanako	roses	here	now.
A3 + T3	+ S3	+ A2	+ A1	+ P1+P2+P3		A3 + P3	+ A2	+ P2	A1 +	P1 +	S3 + T3
主語	時間	空間	間接目的	直接目的	述語部	主語	述語部	間接目的	直接目的	空間	時間

時間と空間に関する単語は日本語の場合は、主語のすぐ後に、<時間+空間>の順で、英語では、文末に<空間+時間>の順に置かれる。主語は両方とも文頭にくる。述語部は、日本語の場合は文末に、意味フレームを構成する基本単文の動詞“P1+P2+P3”がその順に並ぶ。ところが、英語の場合は、基本単文の“主語+動詞”が“A3+P3+A2+P2+A1+P1”のようにならぬ。日本語でも英語でも、P2やP1の動詞は表面に現れず、普通は最下位の動詞P3だけが表現可となるので、このような言語構造であることには気が付かないが、このような構造がときとして顔をだすことがある。基本単文2の主語(花子 Hanako)が間接目的、基本単文1の主語(バラ, roses)が直接目的である。

ここで説明したのは第4文型の“与える”、“give”の場合であるが、日本語も英語も文型と語順は、その意味構造を構成する基本単文と基本単文内での語順によって自動的に決定される。基本単文内での語順は日本語と英語ではほぼ固定されている。したがって、ある動詞の文型はどのような基本単文がどのように組み合わせられているかによって決定される。“浮かぶ”はPS-EX/PS-DO, “与える”はPS-SSS/PS-IS/PS-DOのように、基本単文によって構成されているので、これらの組み合わせで文型や語順が自動的に決定される。これは日本語も英語も同じである。

5. 副詞文

この普通文法では<時間>と<空間>に関する単語は副詞ではなく、時間格(T格)や空間格(S格)という特別席に埋め込まれる単語である。普通言語で副詞となるのは動詞を修飾する単語だけである。副詞文ができるプロセスは使役文や受け身文などの文生成のプロセス(4)とはほぼ同じであるが、副詞文はある文が形容詞の文の主語(A格)に縦に結合し、それがさらに取り出されて、結合した文に逆に結合されるという複雑な言語構造を示す。それを次の例文で説明する。

彼女が歌を歌うのは楽しそうです。 彼女は歌を楽しそうに歌う。
 She sings a song merrily. She merrily sings a song.

この文は日本語も英語も、次の二つの文から構成されている。

文1: 彼女が歌を歌う She sings a song.
 文2: (それ)は楽しそうです (It) is merry.

文1が文2の“(それ)(it)”に縦に結合される。それが、再び取り出されて、今度は文2に縦に結合される。文2が文1と結合する場合、文1を代表する単語はP格に埋め込まれている動詞であるために、文2は文1の動詞に結合する。その様子は次のようになる。

彼女が歌を歌うのは楽しそうです。 She sings a song. : 文1
 ↓は楽しそうですに ↓ is merrily : 文2
 ↓歌う ↓ sings a song : 文1

文1/文2の場合が“彼女が歌を歌うのは楽しそうです”、“She sings a song merrily.”

文1/文2/文1の場合が“彼女は歌を楽しそうに歌う”“She merrily sings a song.”

となる。

6 多様な文の生成

日本語でも英語でもまた他の言語でも自然言語では同じ意味を多種多様に表現することができる。逆に言えば、いろいろに表現されていても、意味が同じであれば同じ内容として人間は理解することができる。何故このようなことが可能であるかは、どのようなメカニズムで多種多様な文が生成されているかが分かれば簡単に理解できる。

6.1 一つの文を二つに分けて、結合させる場合

太郎は明日新潟で花子にバラを与える。

6. 3 その他の多様な表現

文1：本は3冊である

文2：太郎は今日学校で本を買った

を組み合わせると、次に示すようにいろいろな文ができる。

太郎は今日学校で 本が3冊であるの ↓ 本を買った 太郎は今日学校で 本が ↓ 3冊である ↓ 買った

太郎は今日学校で3冊の本を買った。

太郎は今日学校で本を3冊買った。

(太郎は今日学校で本を買ったの) ↓ は (本が3冊であるの) ↓ です。

太郎は今日学校で本を買ったのは3冊であるのです。

(太郎は今日学校で本を買ったの) ↓ は (本が3冊であるの) ↓ です。

太郎は今日学校で買ったのは本が3冊であるのです。

いずれの場合でも，“太郎が今日学校で本を買った”と“その本は3冊である”という文で構成され、その内容はいさかも人間の頭の中では変更されていないので、人間は同じ内容として正確に理解することができる。

7. まとめ

人間の頭の中にある概念構造を記号化したものを普遍言語と名付けた。その普遍言語をベースにして、日本語と英語の普遍的な言語構造と個別言語固有の言語構造を検討した。この結果、以下の点が明らかになった。

1. 日本語や英語などの自然言語の文はいづれも“～がある”，“～である”，“～をする”という3種類の基本単文（基本意味単位）が組み合わせられたものである。
2. どのような複雑な意味を表現する文でもいくつかの文が縦横に結合されたものである。
3. 日本語や英語の単文の語順は、基本単文内での語順によって決定される。
4. 多様に表現される文でも、意味が同じであれば、同じ意味構造を持っている。

参考文献

- (1) Chomsky, N. 1981, Lectures on Government and Binding. Dordrecht: Foris
- (2) チャールズ J. フィルモア, 1988, 格文法の原理, 三省堂,
- (3) 久野すすむ, 1981, 日本文法研究, 大修館書店
- (4) 安藤司文, 1995, 映画で学ぶ知能からみた日本語と英語の普遍文法, スクリーンプレイ出版社
- (5) 安藤司文, 1991, 意味言語の提案, 情報処理学会, 自然言語処理研究会, 情処研報, Vol. 91, No. 80, pp. 73-80
- (6) 安藤司文, 1992, 意味言語から見た日本文の構造—使役文と受け身文, 国語学会平成4年度春季大会要旨, PP. 63-68
- (7) 安藤司文, 1992, 日本語文における論理構造, 国語学会平成4年度秋季大会要旨, PP. 44-49
- (8) 安藤司文, 1992, 自然言語に内在する論理構造, 情報処理学会 自然言語研究会 情処研報, Vol. 92, No. 2, pp. 59-66
- (9) 安藤司文, 1993, 格構造と論理構造との間での変換による多様な文の生成第106回日本語学会大会配布資料
- (10) 安藤司文, 1992, 意味言語から見た主語について, 第104回日本語学会大会配布資料
- (11) 安藤司文, 1991, 意味言語による質問応答, 人工知能学会、ヒューマンインターフェイスと認知モデル研究会、研究資料SIG-III CG-9102, PP. 43-52
- (12) ANDO, S. 1992, The questioning and Answering Based on Meaning Language, International Symposium on Natural Language understanding and AI (ISKIT'92), p. 203-210, Kyushu Institute of Technology, Fukuoka, Japan.
- (13) 安藤司文, 1991, 意味言語による対話と学習, 認知科学学会、学習と対話研究分科会、SIGLAL 91-2, 学習と対話, Vol. 91, No. 2, PP. 1-10
- (14) 安藤司文, 四元健仁, 1992, 意味言語による多言語間機械翻訳, 情報処理学会, 自然言語処理研究会, 情処研報, Vol. 92, No. 55, pp. 49-56
- (15) 安藤司文, 1991, 自然言語に内在する知識獲得のメカニズム, 情報処理学会、データベースシステム研究会、情処研報Vol. 91, No. 84, P
- (16) Ando, S., Masutani, M., 1993, Knowledge Acquisition from Natural Sentences, Natural Language Processing Pacific Rim Symposium Fukuoka, Japan
- (17) 安藤司文, 益谷真, 守屋秀洋, 1993, 多様な自然文からの知識獲得, 情報処理学会 自然言語研究会 情処研報, Vol. 93, No. 101, pp. 25-32
- (18) 安藤司文, 1991, 自然言語に内在する仮説生成機構, 情報処理学会, 人工知能研究会, 情処研報, Vol. 75, No. 80, pp. 1-10
- (19) 安藤司文, 1989, 認知プロセスと自然言語—自然言語の構造(1), 認知科学学会第6回大会 発表論文集, pp. 128-129
- (20) 安藤司文, 1991, 自然言語から意味言語への変換, 情報処理学会第43回全国大会講演論文集(3), pp. 3-137~138