

英語複合名詞句構成の意味関係と意味処理

飯田 仁 小倉 健太郎 野村 浩郷

日本電信電話公社 武蔵野電気通信研究所

1. はじめに

自然言語処理の研究の進展に伴い、日本語、英語にかかわらず構造がかなり複雑な文の意味理解もある程度可能となってきた。その理由の一つに構文、並びに意味に関する辞書内容の充実があげられる。一方、意味理解の程度の高制度化、あるいは処理対象の領域の拡大を目指すとき、その辞書の規模は膨大になる。殊に、語の造語能力を考えると、複合語を逐一辞書に登録することには限界がある。従って、用語を除けば、辞書に登録されていない複合語の扱いは不可欠な処理であり、対象領域を拡大していくとき、その扱い量は増大していく。そこで、未登録複合名詞句の意味処理の頻度が高くなっていく。しかし、複合名詞句は文、又は句が省略表現された抽象的な名詞であり、その句内の意味関係には文内にみられる意味関係と同様、乃至は同等の意味関係が存在する。それ故、複合名詞句は動詞、あるいは特定の述語を用いて句の形にパラフレーズ可能である。そして、その句構成内の意味関係に基づいた意味構造の理解が必要である。また、機械翻訳においても、複合語の訳を語対語の訳とするのではなく、必要に応じパラフレーズを可能とさせるためにも、複合語の意味構造抽出は重要である。

本稿では、複合語のうち複合名詞句に話題を限定して、英語の複合語から眺めた複合名詞句の類別と、複合名詞句内の意味関係の自動抽出法、及びそのために必要な知識と推論法とについて述べる。複合名詞句を類別するに当たっては、ある特定の分類に基づいた複合名詞句内の意味関係の枠内でそれを自動抽出する実験を行い、実験結果をもとに分類の修正、拡張を進めた。また、この実験では、特定分野のテキストから作った用例データベースを参照して意味関係を抽出した。用例との適合の基準は意味マーカー、及びその概念階層体系に基づくマッチングであり、この種のマッチングの効果の程度がこの実験で示された。

2. 英語複合名詞句の研究

複合名詞句を計算機処理するにあたり、複合名詞句内の意味関係を明らかにしておく必要がある。言語学における研究では、J. Levi [1] (Northwestern Univ., 1977) が変形生成文法の枠内で、それらの関係を決定する試みを行った。また、E. V. Clark & H. H. Clark [2] (Stanford Univ., 1979) は名詞が動詞としての機能を果たす事実が多いことに着目し、その動詞と複合語要素の語との意味関係を role としてとらえた。Clark & Clark はそのときの動詞の分類を詳細に行った。計算機処理の立場から見ると、Levi の限定された意味関係の体系を利用していくこ

とが有効である。そこで、まずこの体系が現実をどれ程反映しているか評価した。但し、それら意味関係が有限でないという言語学上の結論にたっている。Levi の体系には無理があることが分かっている。

2.1. 変形文法上の複合名詞句内意味関係

Levi は次の(1)-(3)のパラフレーズにより解釈が異なる点に着目し、複合名詞句内に潜在していると考えられる深層レベルの述語を導入した。

(1) パラフレーズ例-1 (タイプ1)

"a red apple"

=> "an apple which is red"

(記号=>はパラフレーズ可を示す)

(2) パラフレーズ例-2 (タイプ2)

"an apple cake"

=> "a cake which HAVE an apple"

(HAVE; Levi が導入した深層レベルの述語)

=/=> "a cake which is an apple"

(記号=/=>はパラフレーズ不可を示す)

(3) パラフレーズ例-3 (タイプ3)

"apple production"

=> "act of producing apples"

=/=> "production which is an apple"

(1)と、(2)及び(3)との違いは、"which is"を用いたパラフレーズが可能か否かにより判別する。(2)と(3)との違いは、名詞句の主要名詞("headnoun")が動詞の派生語であり、修飾語がその動詞の主格、目的格となるか否かにより判別される。

(1)のパラフレーズが可能な形容詞修飾句は属性と属性値の関係(以下、ATTRと記す)となるが、形容詞修飾句でも、"thermal stress"などはそのような関係ではとらえられない。"thermal"の派生の中心語("center")である"heat"を用いて、この句を次のようにパラフレーズできる。

"stress which is caused by heat"

このパラフレーズは(2)の例であり、この種の"cause"の意味を担う深層レベルの述語"CAUSE"が設定されている。

Levi はタイプ2について9種の述語("Recoverably Deletable Predicate", 以下RDPと記す)を、タイプ3については6種の意味関係(2語以上の修飾句をもつ場合に別に2種を用意)を決定した。そして、タイプ2と、タイプ3を複合名詞句("complex nominals")とした。RDPは深層レベルの述語であるから、RDPの解釈はかなり拡大しておく必要がある。表1にRDP、PNOMの種類と概説・用例を示す。

表1 Leviの複合名詞の意味関係

	意味関係	説明	例
R D P	CAUSE1	修飾語が原因となり主要名詞を引き起こす	malarial mosquitoes, tear gas
	CAUSE2	主要名詞が原因となり修飾語を引き起こす	drug deaths, electric shock
	HAVE1	主要名詞が修飾語を所有する	college town, vegetable soup
	HAVE2	修飾語が主要名詞を所有する	lemon peel, pole eight
	MAKE1	主要名詞が主体となって修飾語を作り出す	musical clock, honey bee
	MAKE2	修飾語が主体となって主要名詞を作り出す	copper coins, student committee
	USE	修飾語が主要名詞の用いる道具となっている	electrical clock, telephonic message
	BE	修飾語が主要名詞と同格的(構造物・属)である	mammalian vertebrates, hairpin turn
	IN	修飾語が主要名詞の実際上若しくは概念上、空間や時間を表す	field mouse, autumnal rains
	FOR	主要名詞が修飾語の為のものとなっている	horse doctor, tax office
P N O M	FROM	修飾語が主要名詞の出どころ(原料・場所など)を表す	olive oil, country visitors
	ABOUT	修飾語が主要名詞の主題を表す	tax law, historical drama
	ACT-SUBJECT	具体的産物を伴わない動詞の派生語で、修飾語がその動詞の主語	staff attempts, student decision
	ACT-OBJECT	具体的産物を伴わない動詞の派生語で、修飾語がその動詞の目的語	musical criticism, child abuse
	PRODUCT-SUBJECT	具体的産物を伴う動詞の派生語で、修飾語がその動詞の主語	royal orders, papal appeals
	PRODUCT-OBJECT	具体的産物を伴う動詞の派生語で、修飾語がその動詞の目的語	tuition subsidies, stage designs
	AGENT-OBJECT	動作主名詞であり、修飾語が派生の中心語である動詞の目的語	city planner, mail sorter
	PATIENT-OBJECT	派生の中心語である動詞の目的語、かつ修飾語がその動詞の主語	royal gifts, city employees

2.2. Leviの体系の評価

2.2.1. 評価の方法

Leviが設定したRDP、PNOMに従って、実際の文章中に出現する複合名詞句内の意味関係をどれ程自動的に決定できるかを調べた。その方法は、用例、又は典型例の知識ベースを作成しておき、文章中より抜き出した複合名詞句を知識ベースと比較し、マッチングをとる。知識ベースとの適合の基準は、意味マーカによる概念の包含関係とした。又、意味概念の体系としては、LEXICON [3] に準拠した体系を用いた。

知識ベースは、LUTEにおける英文解析[4]、[5]、[6]用の名詞辞書に、付加データとして各意味関係で使用される語の意味概念を登録してフレームで構成した。フレーム内の各スロットと、その値の例を図1に示す。図中の数字付きのコードは意味概念を示し、最上位のクラスが25種類、概念階層の深さは3である。

知識ベースには、各語義(例、“energy-noun-1”)に対し、RDP、およびPNOMの意味関係にある用例がフレーム(例、“energy-noun-rdp-1”)で記述される。そのフレーム内には、各意味関係、例えば“IN”、“FROM”がスロットとして扱われ、スロットの値には各意味関係に用例、及び典型例の意味概念が当てられる。

また、Leviが扱っていなかったタイプ1については、10種の属性を設定しRDP、PNOMと同様のフレームによるデータを記入した。属性の全体を表2に示す。

2.2.2. 知識ベースの作成法

汎用の言語データ収集の目的で作った自然言語処理用支援ツールを利用して知識ベースの作成を図った。知識ベース作成のプロセスを図2に示す。

2.2.3. 知識ベースの規模

知識ベースとしての名詞修飾の用例・典型例の収集は、“Scientific American”のエネルギー関連記事の

energy-noun	
スロット	値
type	prototype
pos	noun
root-form	energy
conjugation-s	(ies, y)
meaning	energy-noun-1

energy-noun-1	
スロット	値
type	prototype
category	(H083 H080)
countable	U
RDP	energy-noun-rdp-1
PNOM	energy-noun-pnom-1
ATTR	energy-noun-attr-1

energy-noun-rdp-1	
スロット	値
IN	(L102 L002)
FROM	(L054 H080 C165 L064 H208 H201 H081 H008 O017 L114 L006)
MAKE2	(I073 H008 Y008)
USE	(L063 H008 H201 L006)
FOR	(I073 M090 I100)
CAUSE2	(I075)
HAVE1	(A061)
HAVE2	(L006 H201 L063 L054 L024)

energy-noun-attr-1	
スロット	値
STATE-MODIFICATION	(L024 S007 W002)
GENERAL	(J007)
QUANTITATIVE	(W002 J039 C154)
ACTION-MODIFICATION	(O017)
DEGREE	(R025)
CHARACTER	(L152)
BELONG	(S007)

図1. 名詞辞書の例

6テキスト、約3万5千語より抽出した。出現した名詞の異なり語は、約2300である。そのうち、形容詞、又は名詞による前方修飾語をもつ名詞は711であり、それらについてフレーム記述の名詞辞書を作

表2 実験用に設定した属性の種類

	意味関係	説明	例
A T T R	QUANTITATIVE	修飾語が数量を表す修飾関係	numerous friends, total consumption
	SIZE	修飾語が大きさを表す修飾関係	(a) big man, large units
	SHAPE	修飾語が大小以外の物理的な形状を表す修飾関係	round tables, (a) straight line
	CHARACTER	修飾語が性質を表す修飾関係	beautiful ladies, heavy elements
	TIME-AND-AGE	修飾語が年齢・新古を表す修飾関係	young men, next year
	STATE-MODIFICATION	修飾語が状態を形容している修飾関係	(a) progressive policy, natural body
	ACTION-MODIFICATION	主要名詞が動作を表すか動作を連想する語で、修飾語はその動作を修飾している関係	(a) slow train, radical departure
	BELONG	修飾語が所属を表す修飾関係	human feelings, (a) national center
	GENERAL	修飾語が一般的な性質を表す修飾関係	normal temperature, average automobile
	DEGREE	修飾語が程度を表す修飾関係	low cost, high efficiency

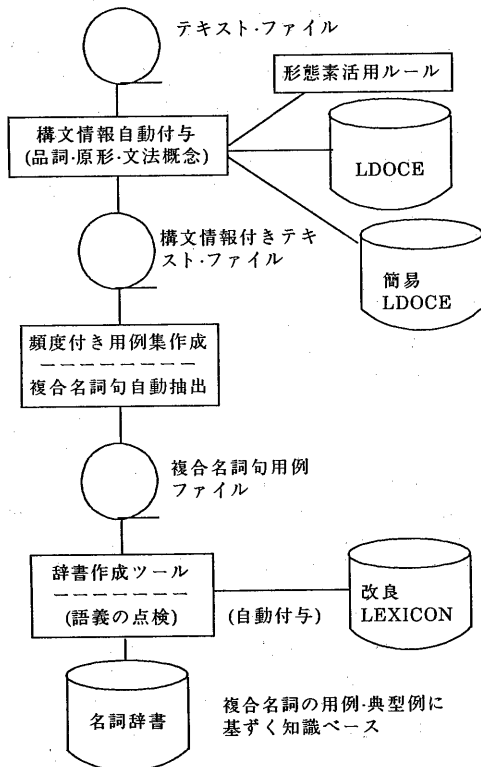


図2. 複合名詞に関する知識ベースの作成プロセス

成した。また、収集対象となった名詞句(以下、“標準名詞句”と呼ぶ)の数は約8200であった。作成に際し、テキスト毎に作られた名詞辞書をマージし、単一エントリーの下にテキスト毎の用例等を統合した。

6テキストは共にエネルギー問題に焦点を当てているが、個々に話題が異なるから、使用語の重複率は50%前後である。但し、同分野についても重複率が35%を下回ると、同一問題を扱っていてもほぼ別の分野とみなした方が良く、分野を限定したものの、用例収集のテキストは6に止まった。

2.3. 意味関係自動抽出の実験結果

知識ベースの作成に使用しなかったエネルギー分野の2テキストから、各2パラグラフ(par1, par2)

を選択し、そこに出現する標準名詞句内の意味関係がどの程度自動的に決定できるかを調べた。

この実験により得られたデータの主な項目は以下の通りである。実験結果については表3.に示す。

- 試料とした標準名詞句の主要名詞と辞書エントリーとの適合率
- 試料の主要名詞の異なりと辞書エントリーとの適合率
- 重複を許した標準名詞句の意味関係の抽出率
- 標準名詞句の異なりに対する意味関係の抽出率
- 名詞辞書中のデータ数を10単位(用例・典型例の総数)以上に限定した場合の(i)-(iv)の値、表3.中に括弧付きの数値で表示

実験結果によれば、テキスト中に出現する重複を許した標準名詞句の意味情報が59%(par1とpar2の平均)抽出できることが分かる。意味関係が適確に抽出できない句は、知識ベース上のデータの不足により生ずる。そこで、知識ベースが充実された状況を想定して、知識ベース上のデータ数を10単位以上に限定した知識ベースに対し同一実験を行った。その結果、81%の意味関係抽出率が得られた。但し、知識ベース上には一意に決定できない意味関係が登録されているので、複数の候補が抽出されることもある。また、使用した意味概念の体系の不備で、常識的に、抽出されるべき関係が導き出せない場合もある。この体系の整備を現在進めているが、問題点、改良点等は稿を改めて述べる。

次に、当実験で意味関係がうまく抽出された句について述べる。

- 概念の包含関係が満たされることによる抽出
例：①hydrogen atomに対するBEの関係(知識ベースにはcarbon atomの例が存在)
②coal research 及び、energy researchに対するABOUTの関係(知識ベースにはbattery researchの例が存在)
- 概念が反意的であることによる抽出
例：①synthetic gas に対するFROMの関係(知識ベースにはnatural gasの例が存在)

2.4. Leviの体系の評価

実験結果よりRDP、及びPNOMの体系を考察すると、計算機処理の立場から主に次の問題点が指摘できる。

表3.意味関係自動抽出の実験

項目	par1	par2
標準名詞句の延べ数:④	87	88
同名詞句の異なり数:⑤	70	71
辞書エントリーが存在した延べ数:⑥	76 (29)	72 (47)
辞書エントリーが存在した異なり数:⑦	59 (24)	56 (33)
(i) (④/④)	87% (33%)	82% (53%)
(ii) (⑤/⑤)	84% (34%)	79% (46%)
④の意味関係抽出数:⑧	41 (23)	45 (39)
⑤の意味関係抽出数:⑨	30 (19)	31 (25)
(iii) (⑧/⑧)	54% (79%)	63% (83%)
(iv) (⑨/⑨)	51% (79%)	55% (76%)

(注)データ数10単位以上を備えた名詞は55語

(1)RDPの解釈がかなり拡張されているので、標準名詞句内の意味関係が多様過ぎる。

(例1) 標準名詞句:“air quality”

意味関係: HAVE2 in RDP、BELONG in ATTR

(例2) 標準名詞句:“energy source”

意味関係: FOR、MAKE1、HAVE2 in RDP

(2)RDPのFORは多くの意味関係に見られるが、主要名詞の性質によってはより明確にその関係が認められる場合がある。

(例3) 標準名詞句:“nuclear plant”

意味関係: FOR、USE

(例4) 標準名詞句:“energy resources”

意味関係: FOR

(3) PNOM の ACT-SUBJECT (OBJECT) と PRODUCT-SUBJECT (OBJECT) との区別が難しい。action、並びに production と言う2種の行為の区別をするより、主要名詞から派生する動詞に基づく格関係を標準名詞句内の意味関係とする方が、関係の明快さ、及び処理の扱い易さの点で有用である。

(例5) 標準名詞句:“power need”

意味関係: PRODUCT-OBJ、ACTION-OBJ

(例6) 標準名詞句:“region need”

意味関係: IN in RDP、PRODUCT-SUB

(例7) 標準名詞句:“price increase”

意味関係: ACTION-SUB

(4)格関係で捕らえるべき他の関係として様態関係を上げることができる。この場合は(3)と異なり、形容詞である修飾語が副詞化して、主要名詞から派生する動詞と様態関係にある。

(例8) 標準名詞句:“suitable development”

意味関係: 抽出関係無し

(例9) 標準名詞句:“direct consumption”

意味関係: GENERAL in ATTR

(5)主要名詞が単位を示し、修飾語が主要名詞に従って特定の形容詞・名詞、あるいは数詞をとる句がある。句全体として単位となる場合が多い。

(例10) 標準名詞句:“cubic foot”、“square

kilometer”

意味関係: SHAPE in ATTR

3. 新しい複合名詞句の体系

前節において変形生成文法に立脚した複合名詞句内の意味関係体系の評価を行った。概して、RDPの述語の定義があいまいであり、抽出される意味関係の解釈の幅が広過ぎる。また、主要名詞、あるいは修飾語の語義に依存して決まる意味関係が多々ある。この意味関係については、計算機処理の立場からみると、領域依存のデータベース検索用に出現する名詞にアプリオリに与えた名詞固有の属性か、あるいは意味関係に対応するとみなせる。データベース・インタフェースである T.Finin 等の JETS [7] (Univ. of Illinois, 1979) では、名詞は“concept”として表現し、今述べた意味的關係をその concept がもつ role としてとらえている。その後、P. Isabelle [8] (Univ. de Montreal, 1984) はそれまでの複合名詞句の研究を再検討し、計算機処理への可能性を探っている。

以下、前節、及び Isabelle の検討を踏まえ、複合名詞句の新しい体系を提示する。また、日本語の漢字熟語にもその体系が適合することを示す。

3.1. 複合名詞句の新体系

3.1.1 機能名詞複合化類

前述した Finin の role は主要名詞から連想される典型的な動詞(逆に、その名詞動詞から派生した名詞)の格関係と捕らえることもできる。同時に、修飾語(名詞、または形容詞の場合はその名詞形)がその格関係で捕らえられれば、主要名詞からの連想動詞に基づく格関係でこの種の複合名詞句の意味関係を捕らえることができる。このような名詞を“機能名詞(role nominal by Finin)”と呼ぶことにする。

Levi の体系の評価で述べた例3の“plant”はこの名詞に当たる。そして、“nuclear”は“plant”より連想される動詞“produce”の対象格となり、plant 自身は場所格となる。その他の、典型例として、

(例11) “bread knife”——— 連想動詞; cut

(例12) “coffee shop”——— 連想動詞; sell

3.1.2. 述語名詞複合化類

名詞が動詞同様に特定の引数をとって、複合名詞句の意味を記述し得る特別な名詞を“述語名詞(root nouns with argument)”とよぶ(P. Isabelle, '84)。この種の名詞は動詞に対応した述語とは表現内容が異なり、この名詞の性質を限定する属性が固定されていて、固有変数を備えた述語表現が可能となる。さらに、動詞からの名詞化による派生名詞も、本来の動詞を述語とする述語表現により複合名詞句の内容を記述できる。但し、今述べた後者の派生名詞は“動詞派生抽象名詞”と呼んで、述語名詞とは独立して扱う(3.1.4. 動詞・形容詞派生名詞の複合化類)。

2.4. 節の例1. に示した“quality”は述語名詞である。その引数として、“対象物”、“質の程度”、“領

域”をとることが考えられる。例1.の“air”は対象物である。その他の、典型例として、

(例13) “base bottom”—— base:対象物

(例14) “room temperature”—— room:場所

3.1.3. 単位名詞複合化類

単位、あるいは単価を示す修飾語とその値を引数にとり得る特別な名詞により構成される複合名詞句がある。これらを“単位名詞複合化類”とよぶ。これらを形成する主要名詞は述語名詞として捕らえられるが、その意味関係は決定的であるので独立した複合名詞句のクラスを設けた。典型例としては、

(例15) “mile race”—— mile:距離

(例16) “nickel cigar”—— nickel:価格

等が上げられる。これらの複合化は述語名詞としての性格が顕著に現れた場合に生ずる。つまり、述語と引数の関係が絶対的な場合であり、この関係が社会常識的に関係づけられることが多いと考えられる。このことは、当複合化類の認識において言語外の知識を援用しなければならないことを示す。但し、2.4節の例10.にみられる句“cubic foot”等は用語として捕らえておくことが適切であると考えられる。

3.1.4. 動詞名詞化複合化類

動詞より派生した名詞から、派生の逆方向に連想される動詞の格構造に従って複合名詞句内の意味関係が定まる場合は多い。これらの句を“動詞名詞化複合化類”と呼ぶ。この種の複合化の出現頻度は一般に高い。典型例としては、

(例17) “energy consumption”

派生の中心語: “consume”、格関係:対象格、

(例18) “manager attempts”

派生の中心語: “attempt”、格関係:行為者格

3.1.5. 形容詞名詞化複合化類

前述の動詞同様に、形容詞より派生した名詞から派生の逆方向に連想される形容詞の格構造に従って複合名詞句内の意味関係が定まる句を“形容詞名詞化複合化類”と呼ぶ。

典型例としては、

(例18) “uranium scarcity”

派生の中心語: “scarce”、格関係:言及対象格、

(例19) “academic freedom”

派生の中心語: “free”、格関係:言及対象格

3.1.6. 動作主名詞複合化類

動作主名詞の派生の中心である動詞の格構造に従って複合名詞句内の意味関係が定まる複合名詞句を“動作主名詞複合化類”と呼ぶことにする。これと、前述の動詞(形容詞)名詞化複合化類と異なる点は

動作主名詞複合化における主要名詞が行為者格、又は道具格になっていて、後者の主要名詞が動作の抽象化となっている点である。又、機能名詞で述べたように動作主名詞は機能名詞としての役割を果たしているが、派生の中心である動詞を名詞の形態から決定することが出来る点で当複合化類を機能名詞複合化類から独立させた。つまり、意味処理上、動詞の連想過程を明確に切りわけることが可能となる。典型例としては、

(例20) “truck driver”

派生の中心語: “drive”、

対象格: “truck”、行為者格: “driver”

(例21) “mail sorter”

派生の中心語: “sort”、

対象格: “mail”、道具格: “sorter”

3.2. 特殊な形容詞修飾類

形容詞修飾の多くの標準名詞句は2.1節で述べたタイプ1に分類されるが、形容詞の中にはタイプ1のパラフレーズを行うと形容詞の意味が異なってしまうものがある。また、形容詞修飾でもその属性が限定されることがほとんどなく、単に主要名詞の意味が強調されているだけと認められる場合がある。以下に、これら2種の場合について述べる。

3.2.1. 特殊な限定修飾類

タイプ1のパラフレーズを行うと形容詞の意味が異なってしまう形容詞は、それ自身が主要名詞の属性値を決定してしまう。例えば、“late president”において“president”の属性として、“任期期間”等を設定しても属性の十分な分類に役立つものではない。特に計算機処理を考慮すれば、“役職”や“地位”に対して特殊な属性を仮定して設定しておくことが考えられる。今の例であれば、属性“LATE-attribute”に対して正か否かの指標として“T/NIL”を取ることが分かれば良い。他の典型例として

(例22) “present government”

を上げることができる。しかし、これらの形容詞の数は極わずかであることが一般に知られている。その例は、以上の他に“glad”、“poor”である。“certain”もこの種の語であるが、これは次節3.2.2.の複合化類で扱う。また、タイプ1のパラフレーズの結果にみられる形容詞の限定的用法はすべての形容詞が一般に備えている性質ではない。物質名詞に“en”が結合した“物質形容詞”や“既往比較級”は、多くの場合限定的用法しかとらないとされている[9]。

3.2.2. 強意修飾類

主要名詞の属性を限定することがない修飾語を備えた標準名詞句を“強意複合化類”と呼ぶ。

典型例としては、

(例23) “certain place”

(例24) “total nonsense”

を上げることができる。その他、この種の修飾語としては、“mere”、“sheer”等である。また、“white snow”の如く主要名詞が本来備えているべき属性値(デフォルト値として)を誇張する修飾用法があるが、現段階ではこれを通常の形容詞修飾用法として扱っておく。

3.3. 標準名詞句構成の意味関係の新体系

3.1.節で述べた複合名詞句の意味関係は特徴的な名詞の性質により決まるが、関係の決まり方にはもう少しのヴァリエーションがある。このヴァリエーションに触れる前に、まず標準名詞句構成の意味関係について大分類をしておく。3.1.節、及び3.2.節の検討から、五つの大分類が得られる。それらは、

(I)機能名詞複合化類、及び述語潜在化類に代表される複合名詞句の集合;

これらは動詞、あるいは潜在述語を介して修飾語と主要名詞との関係付けがなされる。

(II)述語名詞複合化類、及び単位名詞複合化類の集合;
これらは主要名詞が述語的な役割をはたしている、修飾語との関係がその述語により決定される。

(III)強意表現を除いた形容詞修飾類の集合;

修飾語と主要名詞の関係により主要名詞の属性が限定される。

(IV)強意複合化類の集合;

修飾語により主要名詞の属性が限定されない。

(V)動詞、並びに形容詞の名詞化複合化類の集合;

これらは動詞・形容詞の格構造で意味関係が捕えられる。但し、格関係である様態関係については、英語の場合、修飾語の形容詞から副詞への派生により様態関係が導出されるので、上述の二種の名詞化複合化類とは別のクラスを設定する。

大分類について補足すれば、クラス-Iでは動作主名詞もそれより連想される動詞に基づき意味関係が決まるが、形態素のレベルでその動詞の決定ができるので、これらの複合名詞句を“動作主名詞複合化類”と呼んで独立させる。さらに、連想される動詞が直接的に確定せず、何段かの推論を必要とする機能名詞的な複合化も存在する。例えば、

(例25) “soap heir”

・“heir”とその基礎概念(意味概念体系の基礎概念として“Legal-work”)より連想される述語は“inherit”であるが、今、このinheritの対象格に話しを限定して進めると、その対象格として典型的に導出できる語の一つは“fortune”である。

・“fortune”も機能名詞としてとらえ、そこから連想される述語として“make”(“財産を築く”と言う意)を得る。そして“make”の対象格である“fortune”と源泉格となる“soap”との関係が決定

dekiru。このとき、“soap”を源泉格の要素と見なすためには言語外の社会的な知識を必要とする。

この種の言語外の社会的な知識を用いて意味関係が導出されるクラス-Iの複合名詞句を“機能名詞依存構造化類”と呼ぶ。

次に、クラス-Vで触れた様態関係をもつ複合名詞句は“様態関係複合化類”と呼び、以下にその例を示す。

(例26) “technical innovation”

・“innovation”より連想される述語:“innovate”

・“technical”より連想される様態関係:

technically

(例27) “physical collapse”

(例28) “total stranger”

この場合は、“stranger”より連想される述語を便宜的に“BE-STRANGE”とし、動作主名詞複合化類として捕えられるが、修飾語との関係がはっきりしていることからこの複合名詞句を様態関係複合化類に含める。

RDPは詳細に分類する必要があるので、LUTEにおける英文の格解析時と同様に潜在述語の語義に応じて、RDPの詳細化を進めた。特に、BEは多くの意味をもつが、“child professor”([1])の様にBEの近似的な述語として“BE-SIMILAR-TO”等を用意したほうが良い場合がある。

以上述べた複合名詞句を含む標準名詞句の新しい体系を付録1に示す。

3.4. 日本語との対照

日本語の漢字熟語の構成について調べると、野村雅昭[10]の体言類、用言類、相言類、副言類に基づいた分類、影山太郎[11]の日・英語の語形成に関する考察、田中康仁等[12]のバリエーションを中心にした分類等がある。これに対し、前節までに述べてきた新体系では格構造による意味関係の把握にとどまらず、名詞が潜在的に備えている複合化の機能等を明らかにしてきた。そして、この体系は日本語の漢字熟語の構成にも適応可能である。詳細は避けるが、以下に主要部の対比と日本語の複合名詞句例を示す。

日本語において多用される<目的語>+<用言類>形、あるいは<相言類>/<副言類>+<用言類>形はクラス-Vに対応しているが、<用言類>+<用言類>形(“立ち読み”等)は英語では等位接続詞を用いた表現になるので、新体系からは除外されている。他に、<体言>+<体言>(“男女”)、<相言類>+<相言類>(“甘酸っぱい”)等の一部は同様な理由で新体系からは除外されるが、<体言>+<体言>は述語潜在化類、もしくは機能名詞複合化類等のクラス-Iで扱える。また、RDPの述語を詳細化したほうが良い例としてあげた“BE-SIMILAR-TO”についても、“こぬか雨”の様な日本語例がある。

クラス-IIIの特殊限定修飾複合化類の日本語との対応をみると、“前”、あるいは“現”等の接頭辞による複合化がそれに対応している。また、クラス-

IVの強意複合化類は“真”等の接頭辞による複合化類に対応する。ただし、日・英語ともに文脈によっては単なる強意ではなく、意味を持つ場合がでてくる(“Aは犯人だ”に対する“Bは真犯人だ”においてみられる“真”の意味)。

以下の表4に、標準名詞句の構成にみられる意味関係に基づいた12種の分類体系と日本語の用例を示す。

表4. 標準名詞句構成の分類と日本語の例

大分類	標準名詞句構成の分類	日本語の例
I	機能名詞複合化類	刺身包丁
	動作主名詞複合化類	会社経営者・水先案内人
	述語潜在化類	マラリア熱・税法
	機能名詞依存構造化類	土地成金
II	述語名詞複合化類	室内温度(室温)
	単位複合化類	円タク・100m競争
III	限定修飾複合化類	丸顔・特殊兵器
	特殊限定修飾複合化類	(前……, 現……に対応)
IV	強意複合化類	真水・真犯人
V	動詞名詞化複合化類	ねじ回し・雨上がり
	形容詞名詞化複合化類	物価高・品行方正
	様態関係複合化類	早起き・急上昇

4. 新体系の意味関係の処理と知識

4.1. 意味処理に必要な基本知識

標準名詞句内の意味関係の自動抽出、あるいは意味関係に基づく標準名詞句の自動収集を進めるとき、まず必要となる基本的な知識を明らかにする。

(1) 形態的な知識として、形容詞からその派生の中心である名詞を導出するための知識。

例えば、“thermal”から“heat”を導出するために必要となる知識である。

(2) (1)と同じく、名詞からその派生の中心である動詞、あるいは形容詞を導出するための知識。

例えば、“consumption”から動詞、“consume”を、また“freedom”から形容詞、“free”を導出するための知識である。

(3) 前二者と同様、形態的な知識として、形容詞から副詞化を行うために必要な知識。

例えば、“physical”から“physically”を導く知識である。

(4) もう一つの形態的な知識である動作主名詞の動詞導出のための知識。

例えば、“driver”に対する“drive”という知識である。

(5) 次は、形態的なレベルでは導くことができない知識であり、機能名詞としての役割を果たせるか否かを判断する知識。

例えば、“shop”から“sell”という述語を連想するための知識がこれに当たる。この種の知識は言語外の知識が反映されているものの、各語が本来備えているべき意味に依存すると考えられるので、この様な連想知識は言語内の知識と考える。

(6) 機能名詞依存構造化類における連想述語導出のための言語外の知識。

例25.にみられるように、この種の知識は典型例としての連想関係を明記することで表記できる。

(7) 述語名詞複合化類を作る述語名詞であるか否かの判定と述語に関する知識。

例えば、“bottom”に対して

Bottom-述語(変数-対象物 変数-内容物)

なる述語を導出する知識

(8) 特殊限定修飾複合化類を作る特殊な形容詞であるか否かを判定する知識。

これらのクラスは前述した{glad, late, poor, present}である。

(9) 単位複合化類に係わる知識は“race”に見られるように、

Race-述語(変数-距離)

なる述語を導出すると同時に、修飾語が単位という概念下にあることを判断する知識。

(10) 知識ベースとして辞書に備えておく知識としては、各連想述語の格フレームと名詞の意味マーカ、及び意味素性である。

以上が意味関係の処理に必要な基本知識である。従って、機能名詞依存構造化類における複数段の推論においても、上述の基本知識を繰り返し使用することで意味関係の導出が可能となる。

4.2. 知識ベースの表現

前節で述べた基本知識は図1.でみた名詞辞書のスロット中に各連想知識が記述され、格解析に必要な動詞、あるいは述語にアクセスできる。また、図1.で示したRDP等の各スロットの値(用例を満たすための意味マーカ)に対応して、制約条件をフレームで記述し、条件スロットに次のタイプの条件をかく。

(AND (semantic-category C101); 改良LEXICON
(semantic-feature +concrete)
(syntactic-feature +THE_or+POSS_))

先の2.節で述べた実験では、知識ベースに記述された用例データ等は主要名詞に関して辞書記述されているが、クラス-IIIでは形容詞が中心的な役割を果たしているため、形容詞側にどのような名詞を修飾するかを記した知識を記述しておく。図1.のフレームに沿った例をあげる(図3.)。

4.3. 意味処理

意味処理における基本操作は各辞書を参照しつつ、連想述語に基づいて格解析を進める。また、クラス-IIIについては形容詞側から用例、並びに典型例として書かれた知識ベースを参照して属性/属性値等の関係を定める。しかし、機能名詞等への連想が可能な名詞であっても、それが機能名詞複合化類に必ず含まれるとは限らない。“knife”に対しての“good knife”や“plastic knife”などはそのような例である。すなわち、付録1.の中央部の枠内は対象とする標準名詞句が総ての複合化類に落ちる可能性を示していて、意味関係決定の為には原則的にはあらゆる可能性を調べる必要がある。但し、各名詞、及び

round-adjective-1	
スロット	値
ATTR	round-adjective-ATTR-1
(以下省略)	...

round-adjective-ATTR-1	
スロット	値
attribute	shape
requirement	(and (modifier-semantic-cat D112)
	(or (semantic-feature +solid)
	(syntactic-feature +animate)))
(以下省略)	...

図3. 形容詞語義辞書の例

形容詞が複合化においてどのような性質を有するか、知識ベースを参照しつつ無意味な処理をしないよう進める必要がある。

5. まとめ

標準名詞句構成の意味関係を自動抽出する実験を行い、それに基づいて新しい複合名詞句、並びに限定修飾句の体系を構成した。同時に、この実験では意味マーカを用いた意味処理の可能性として、十分な辞書内容が用意されていれば、80%前後の自動抽出率が得られる事を示した。また、ここで述べた複合名詞句等の新体系は英語ばかりでなく、日本語にも適用できる事を示した。

参考文献:

- [1] Levi, Judith N., "The Syntax and Semantics of Complex Nominals", Academic Press, 1978
- [2] Clark, Eve V. and Clark, Herbert H., "When Nouns Surface As Verbs", Language, Vol.55, No4, pp.767-811, 1979
- [3] MacArthur, Tom, "Longman LEXICON of Contemporary English", Longman, 1981
- [4] 飯田 仁, 小倉 健太郎, 野村 浩郷, "ATNと格解析を融合した英文名詞句解析", 情報処理学会自然言語処理研究会資料34-8, 1982
- [5] 飯田 仁, 小倉 健太郎, 野村 浩郷, "機械翻訳システムLUTEにおける英文解析", 情報処理学会自然言語処理研究会資料35-4, 1983
- [6] Iida, H., Ogura, K. and Nomura, H., "A Case Analysis Method Cooperating with ATNG and Its application to Machine Translation", COLING-84, pp.154-158, 1984
- [7] Finin, T., Goodman, B. and Tennant, H., "JETS: Achieving Completeness Through Coverage and Closure", IJCAI, pp.275-281, 1979
- [8] Isabelle, Pierre, "Another Look at Nominal Compounds", COLING-84, pp.509-516, 1984
- [9] 大塚 高信, 他監修, "新英語学辞典", 研究社, 1982
- [10] 野村 雅昭, "造語法", 岩波講座日本語9, 語彙と意味, 岩波書店, 1977
- [11] 影山 太郎, "日英語の語形成", "講座日本語学12, 外国語との対照III", 明治書院, 1979
- [12] 田中 康仁, 水谷 静夫, 吉田 将, "語と語の関係について", 情報処理学会自然言語処理研究会資料41-4, 1984

付録1

