

日本語助数詞の分析

— 名詞と助数詞の統語的差異 —

亀井 真一郎 村木 一至

NEC 情報メディア研究所

1 はじめに

本論文では日本語の助数詞を分析し、名詞と助数詞の統語上の区別を明確化する。日本語の数量表現に関しては従来から多くの研究がある。言語学の分野の研究 [1, 2, 3, 4] では数量句の一文種での移動を中心に分析が行われてきた。機械翻訳における処理等に関する言語工学的研究 [5, 6, 7, 8, 10, 11] においては、数量句の文節レベルの構成の定式化と、適切な助数詞の選択を中心に研究がなされてきた。しかしながら、実際のテキスト処理を行なうにあたって個々の単語の辞書を作成する際には、日本語の助数詞の品詞としての定義が問題になる。

日本語において名詞の数量を表現するには例えば「3本のペン」の「本」のように必ず助数詞が用られるのが原則である。日本語では「3ペン」のように数詞を名詞に直接前接させて名詞の数量を表現することは原則として出来ない。しかしながら「袋」という名詞は「一袋の米」のように助数詞と同様な計量子として用いられることがあり、その際には数詞と直接接続する。また「語」という名詞はそれ自身の数量を表すのに「一語」のように数詞を直接前接させることが出来る。後者の語群を吉武ら [8] は NumN 型名詞と呼んだが、それらと通常の名詞との違いが何に起因するかについては明らかにされていない。また前者の「袋」ような名詞がどのような語類であるかについても解明されていない。

助数詞 (接辞) と名詞 (自立語) とは明らかに異なる品詞であるにも関わらず、両方の用法を持つ上記のような語群が存在することから、名詞と助数詞の区別は曖昧になる。このような事情により従来は実際の辞書構築や文法記述の際にしばしば助数詞と名詞の取り扱いがルーズになり、名詞句解析の品質低下の一因となっていた。

そこで我々は数詞に直接後接する語を網羅的に収集し、それらを意味的、統語的に分類した。その結果、助数詞が6つの意味類に分類できることが明らかになった [9]。本論文ではこの意味類の観点から名詞と助数詞の品詞設定を明確化する。さらにこの意味類の観点から見ると、従来全く異なると考えられていた英語と日本語における名詞の数量の表現の仕方が共通の意味的枠組に基づいていることを指摘する。

なお日本語の助数詞には数詞に前接するもの (例: 約) と後接するもの (例: 枚) とがあるが、本論文では後者のみを分析の対象とする。また、数詞に後接する助

数詞には大別して数量の大きさを表現するもの (例: つ) と順序を表現するもの (例: 番目) との二種類があるが、本論文では前者のみを対象にして説明する。

2 名詞と助数詞の差に関する仮定とその問題

以下の (1) ~ (3) が日本語における名詞の数量を表現する際の基本構文である。なお通常は、これらの数量表現構文において数詞に後接して名詞を計量する語類を助数詞と呼ぶことが多いが、本論文では「計量子」と呼んで語の品詞としての助数詞と区別する。

(1) 数 計量子 の 名詞 (助詞)

ex. 一 つ の 時計 (を)

(2) 名詞 数 計量子 (助詞)

ex. 時計 一 つ (を)

(3) 名詞 (助詞) 数 計量子

ex. 時計 (を) 一 つ

いずれの構文の場合にも、名詞の数量を表現するには「つ」のような「計量子」が必要である。名詞に直接数詞を前接させて例えば「一時計」のように表現することは通常できない。このことから、名詞と助数詞は、数詞と隣接するか否かの観点からは以下のように区別されると仮定できそうである。

[仮定: 名詞と助数詞の差異]

名詞 … 数詞が直接前接することはない

助数詞 … 数詞が直接前接する

しかし実際に個々の語の振舞いを見るとこの仮定には問題があることがわかる。数量表現構文の「計量子」の位置に用いられる語群としては実際には様々な種類がある。まず第一に以下の語類Aがある。これらは必ず数詞に後置して使われるかまたは「本数」「枚数」のような特定の構文中にのみ使われ、単独で自立語として用いられることがない。この点でこの語類は典型的な助数詞であり、名詞とは明確に区別される。

語類A: 典型的な助数詞

ex. つ、本、枚、冊

第二に、通常の普通名詞でありながら、自分自身の数量を数える時に数詞が直接前接しうる語類Bがある。

語類B：数える時に数詞が直接前接しうる名詞
ex. 語、票、行、列

第三に、数量表現構文において数詞に後置して計量子として使われるだけでなく、単独で名詞としても使われる語類Cがある。

語類C：計量子として用いられうる名詞
ex. 組、杯、袋、種類、回

典型的な助数詞すなわち語類Aを考えている限り、助数詞と名詞の差異は数詞が直接前接するか否かという区別でよいが、語類Bのや語類Cが存在することから、数詞が前接するか否かという名詞と助数詞の区別はさらに詳細に検討する必要があることがわかる。

3 計量子の意味類と名詞の振舞い

3.1 計量子の意味類

そこで我々は数量表現構文の「計量子」の位置に表れる語、すなわち数詞に直接後接し得る語を網羅的に収集し、それらを構文的、意味的に分類した。その結果、数量表現構文の「計量子」は表1のような意味類に分類できることがわかった。表1に示したように、計量子が表す「多さ」は次の4種類に分けられる。

1. ものの集まりの全体量
2. ものの種類の数 (ex. 種類、通り)
3. もののもつ属性量 (ex. cm、才)
4. 行為の頻度 (ex. 回、周り、循)

さらに「ものの集まりの全体量」の数え方は次の3種類にまとめられた。

- 1-1. それ自体を一つ一つ数える (ex. 人、個、匹、枚)
- 1-2. 組にまとめて数える (ex. 組、足、段、層)
- 1-3. 分量器を使って数える (ex. 杯、袋、箱、皿)

すなわち計量子は 1-1. 1-2. 1-3. 2. 3. 4. の6種類の意味類に分類できることがわかった。本論文ではそれら6種類の計量子の意味類をそれぞれ PIECE、SET、CONTAINER、KIND、UNIT、TIMES と呼ぶ。

なお数える名詞の意味や様子によって使われる計量子に制限があることはよく知られている。例えば、1-1.

何をどう数えるか		計量子の種類
1. 全体量を	1-1. 一つ一つ	(A) 数 計量子 [+PIECE] の名詞 - 中立 - 物の形状 - 名詞の意味
		(B) 数 名詞 [+Element]
	1-2. 組に まとめて	数 計量子 [+SET] の名詞 - 中立 - 構成要素数 - 集まりの形状 - 名詞の意味
	1-3. 分量器を 使って	数 計量子 [+CONTAINER] の名詞
2. 種類を		数 計量子 [+KIND] の名詞
3. 名詞の属性量を		数 計量子 [+UNIT] の名詞
4. 行為の頻度を		数 計量子 [+TIMES] 動詞する - 中立 - 行為の形状

表 1: 計量子の意味分類

の計量子には次のような類がある。

- ・中立的 (ex. つ)
- ・名詞の形状を含意 (ex. 本、片、枚、粒)
- ・名詞の意味分類に依存 (ex. 人、冊、台、脚)

同様に 1-2. の助数詞には以下のような類がある。

- ・中立的 (ex. 組、セット)
- ・構成要素数 (ex. 対、ペア)
- ・構成要素数+名詞の意味分類 (ex. 足、つがい)
- ・集まり方の形状 (ex. 重、段、層、束)

3.2 名詞の振舞い

前述のように名詞の中は数える時に数詞が直接前接しうるもの (語類B)、および計量子としても用いられるもの (語類C) がある。前節でみた計量子の意味類の観点から分析すると、どのような名詞がこれらの語類に属するのかが明らかになる。

まず語類Bについて考える。この語類には、その名詞自身がさらに大きなものの構成要素であるという意味的特徴がある。例えば次のような語が語類Bの例である。

語類B：	(例)
それ自身が素単位	点、ポイント、株、票
階層、序列の要素	並び：行、列
	語：文字、語、文節、文、段落
	鉄道：駅

表1では1-1.(B)に名詞 [+Element] として語類Bを表した。すなわち、名詞を一つ一つ数える場合、通常は1-1.(A)のように計量子が必要であるが、語類Bすなわち [+Element] の意味素性をもつ名詞の場合には、「計量子 [PIECE] + の」の部分を省略して直接数量と名詞とが接続しうる。

なお語類Bの名詞に類するものとして以下のような名詞的接辞がある。これらは単独で自立語として機能しないが、数詞と結び付いたり漢語系指示詞と結び付いて名詞を形成する。これらには人あるいは組織を表すという意味的特徴がある。

紙、誌、社、行、校
(三紙、数紙；各誌、全誌、本誌；貴社、弊社)

次に、計量子としても用いられる名詞類 (語類C) について考える。語類Cの意味的特徴と表1に示した計

量子の6つの意味類を比較した結果、これらの計量子の意味類に属する意味をもつ名詞は、一般に計量子としても用いられることがわかった。例えば名詞「袋」「皿」は語類Cであるが、これは何かの全体量を計る時に分量器として用いられ得る意味的性質 ([+CONTAINER]) をもつ。

我々が収集した語に関する限り、名詞であって数詞に直接隣接するものは全て上記の語類Bまたは語類Cに分類できた。その他の一般の名詞が数詞に直接後接することは通常ない。仮に通常の名詞が数詞と直接接して用いられた場合には、文脈上、語類B (さらに大きなものの構成要素である) または語類C (計量子の6種のいずれかの意味類である) と解釈されて使われていると考えられる。例えば、空間の広さを表現するのに「部屋」を分量器として用いた場合を考える。この時「部屋2つ分の広さ」と表現してもよいが、これを「2部屋の広さ」と表現したとすると、この表現における「部屋」はCONTAINERあるいはUNITとしての意味役割を担っている。CONTAINERの類の計量子が固定的でなく生産性があるのは、文脈上、CONTAINERの意味役割を担い得る名詞が多く存在するからである。

3.3 品詞の設定と辞書構成

本節では前節までの分析に基づき、従来品詞が曖昧になりがちだった語類B、語類Cの辞書構成について考える。まず典型的な助数詞 (語類A：ex. 個、枚) は以下の特徴をもつ。

- ・単独で自立語として用いられない。
- ・数詞と共に用いられた時、その数詞は助数詞ではなく名詞の量を表す

これに対して語類B (例：語、文) は以下の特徴をもつ。

- ・単独で自立語として用いられる。
- ・数詞と共に用いられた時、その数詞は自分自身の量を表す。

この点で語類Bは語としては典型的な助数詞とは異なる。しかしながら「数詞+語類Bの語」からなる数量句は「数詞+助数詞」からなる数量句と同様の数量表現構文をとりうる。従って語類Bの品詞としては、(1) 語としては自立語 (名詞) である、(2) 数詞と隣接する性質があり、助数詞と同等な数量句を形成する、という性質を持つ名詞の下位分類を設定するのが妥当である。

次に語類Cを考える。この語類は単独で自立語として使われる場合と、計量子として使われる場合があるが、その二つの場合で何を数えているかに違いがある。

- (i) 2 つの袋
- (ii) 2 袋のじゃがいも

上記(i)では「袋」自身が「つ」という計量子によって数えられている。これは通常の名詞の構文である。これに対し上記(ii)で数えられている対象は袋自身ではなく、後ろにある名詞(この場合「じゃがいも」)である。名詞が陽に表現されず数量句が単独で「2袋食べた」のように用いられた場合を考えればそのことは明らかである。これは対応する英語が通常とは異なる係り受けをもつ名詞句構文「two bags of potatoes」をとることに反映している。「何を数えているか」という観点での意味の違いは上記(i)(ii)の構文の違いに対応しているため、語類Cに関しては、辞書上の品詞として上記(i)に対応する名詞と上記(ii)に対応する助数詞とを設定すると取扱上都合が良い。

4 日本語と英語の比較

本節では、表1に示した計量子の意味分類の観点から、日本語と英語の比較を行なう。通常、日本語と英語とでは名詞の数量の表し方に大きな違いがあると考えられている。英語では通常「two men」のように数詞が直接名詞を修飾するのに対し、日本語では通常「二人の男」のように計量子が必要であるからである。

しかしながら、表1に示した計量子の分布において日本語と英語とで違いがあるのは1-1。「全体を一つ一つ数える」場合だけであることがわかる。他の5つの意味類の場合には、英語でも計量子(counter)が用いられる。以下はその例である。

+SET	2組(の～)	two sets (of～)
+CONTAINER	2袋(の～)	two bags (of～)
+KIND	2種類(の～)	two kinds (of～)
+TIMES	2回(の～)	two times (of～)

英語と日本語とで違いのある1-1の意味類は、名詞の可算/不可算と関連づけることができる。英語の名詞には可算名詞(例: a man)と不可算名詞(a piece of furniture)との区別があり、名詞の数量を表現する時、前者は数詞を直接前に伴い、後者は計量子(counter)を介す。この可算/不可算(数量を表現する時に計量子を伴うか否か)の観点から日本語の名詞を見ると、大部分の日本語の名詞は不可算名詞に対応し、語類B(例: 語、文)が可算名詞に対応すると言える。

以上のように、従来大きく異なると考えられてきた日本語と英語の名詞計量の構造は、表1で示される同じ意味的枠組に基づいていると言うことができる。

5 おわりに

本論文ではまず計量子を意味類が6種類にまとめられることを示した。そしてこの意味類に属す意味をもつ普通名詞は、自身が助数詞として働き、他の名詞を数える機能を持ち得ることを述べた。また普通名詞であっても意味的に[+Element]の属性を持つ語は、自分自身を数える時に助数詞を介さず数詞を直接前接させ得ることがわかった。これらの分析に基づいて従来曖昧になりがちであった助数詞と名詞の品詞設定の基準を示した。

また従来は、日本語と英語の名詞の数え方は大きく異なるものと考えられてきたが、この意味分類に基づけば、その差は[+Element]の部分だけであることを示した。日本語の中にも不可算名詞と可算名詞の区別を見出すことができることを述べた。

今後は、助数詞の発達している他の言語、例えば中国語[11]や韓国語を分析し、本論文でのべた計量子の意味類の枠組をさらに整備してゆきたい。

参考文献

- [1] 神尾昭雄 1977. 「数量詞のシンタックス」月刊言語 Vol.6, No.9, pp.83-91.
- [2] 井上和子 1978. 「日英対照 日本語の文法規則」大修館書店
- [3] 柴谷方良 1978. 「日本語の分析」大修館書店
- [4] 奥津敬一郎 1983. 「数量詞移動再論」人文学報 160号 pp.1-24. 東京都立大学人文学部
- [5] 田中康仁 1980. 「日本語の数表現の解析」情報処理学会計算言語学研究会 64-4
- [6] 天野、河田 1980. 「かな漢字変換における局所意味解析」昭和55年度電子通信学会全国大会
- [7] 飯野、亀井 1989. 「英語数量表現の解析」第38回情報処理学会全国大会論文集 3 pp.394-395.
- [8] 吉武、高山、鈴木、太細 1989. 「日本語の数量表現とその英語への機械翻訳に関する一考察」情報処理学会 自然言語処理研究会 74-7.
- [9] 亀井、村木 1990. 「日本語助数詞の分析」第41回情報処理学会全国大会論文集 3 pp.155-156.
- [10] 旭岡、平川、天野 1990. 「日本語の数量的表現の意味処理について」情報処理学会 自然言語処理研究会 78-17.
- [11] 任福継 1992. 「JCMT システムにおける訳文の特殊な数量表現に関する一考察」情報処理学会 自然言語処理研究会 91-1