

# 辞書・事典とは何か

—その体系性・網羅性・信頼性を求めて—

長尾 真

国立国会図書館長

# 辞書・事典の区別

- 辞書は言語の構成要素である語とその説明を集めたもの
- 事典はそれ以外の各分野の対象についての説明を集めたもの
- いずれの場合も人間のもつ概念とその表現とについて記述したもの
- 現実問題として辞書と事典の区別を明確にすることは難しい

# 辞書学 (lexicography)

- 言語の構成要素である語をどの範囲で集め、どのように配列するかが辞書学の主要課題である。  
あいうえお順、漢字順、逆引き、シソーラス配列、オントロジー配列、……
- 語の単位をどのように選ぶか (語構成論)  
できるだけ短い単位にする  
高い頻度で使われる複合表現を単位とする  
できるだけ長くとり  
句単位にとる

- 語の記述項目をどのように選び、どのように配列するか

見出し語(カナ表記)、発音、シラブル区分

同義語、反対語、類義語、……

意味内容の説明(内包的定義)

語の用例(外延的説明)、慣用句

派生語、合成語、……

# 語い論 (lexicology)

- 語の説明を種々の観点からするために各語について研究する分野
- 語源、語の意味の時代的変遷、語の意味内容(多義の場合は別々に)の説明
- 語の用法を種々の異なった例などで示し、その用法の説明をする
- 語の誕生、他の語との関係、慣用的な使われ方(他の語との共起関係)などの説明

# 辞書の例

- 単語辞書、複合語辞書、類義語辞書、漢字辞書、
- 逆引き辞書、語源辞典
- 慣用句辞書、用例辞書、……
- 英和・和英辞書、……
- 辞書と事典の機能を1つにまとめた辞典  
新世紀ビジュアル大辞典(学研)  
Larousse Dictionaries にも良いのがある。

# 事典 (encyclopedia) について

- 百科事典、医学事典など、各分野の事物について説明したもの。
- 見出し項目のたて方によって大項目主義と小項目主義とがある。

Encyclopedia Britanica (大項目主義)

Encyclopedia Americana (小項目主義)

- 事典の場合、どの範囲をおおうかを明確にすること、およびその範囲に存在する概念をどのような概念体系でとらえ、それぞれの概念をどの大きさの単位で取り出すか、即ち概念の粒度をどのように設定するかが問題となる。
- また1つの分野の1つの概念がその分野の概念体系の中のどこに位置づけられるかを明示することが大切である。  
(オントロジーの構築)

- 岩波情報科学辞典 (Encyclopedic Dictionary of Computer Science, EDCS) では、小項目主義の立場をとりながら、用語の木 (オントロジー) を導入することによって大項目主義的にも使えるという方法を案出した。

# 見出し項目の粒度

- 1つの分野には種々の概念が存在し、それはオントロジーの体系をなしている。
- その分野に存在する全ての用語を集め、これをオントロジーの体系(用語の木)に編成する。
- 用語の木の末端の細かい粒度の概念の用語をどこまで事典の見出し語として採用するかを検討する。

- 事典の見出し語の粒度をある程度そろえる。
- EDCSでは1.5万語の用語を集め、用語の木を作り、見出し語として約4500語を採り、その他の粒度の小さい8500語は見出し語の内容説明のところで理解できるようにした。人名その他の固有名詞は粒度の小さい語とみなされるが、利用者の関心の高い語は採用した。

# 用語の説明

- 同義語、反義語、上位・下位概念語、関連語を示し、その語のその分野における位置づけを示す。
- その用語の意味を過不足のないように記述する(内包的定義)。
- その用語のできるだけ異なった多くの事例を与える(外延的定義)。

- その語について、上位・下位関係のほかに、部分全体関係、性質・属性・機能関係、目的関係、因果・継承関係などの記述を行う。
- 用語の説明において大切なことは、類似する用語や関連する用語との関係、その差異を明確にすることである。
- 共時的説明のほかに通時的説明（歴史的な流れの中での記述）を行う。

# 用語間の関係性ネットワーク

- 用語の説明で示された用語間の関係をたどってゆけるように事典構成を行う。
- 概念体系から個別概念へ、またそのルートを作る。
- 個別概念から関連する概念へ(上位・下位関係、部分・全体関係、性質・属性・機能関係、目的関係、因果・継承関係など)、またそのルートを作る。

# 索引と検索

- 専門分野の用語は多くは長い複合語となっている。
- 長い複合語を正確に記憶していることは少ない。  
国連気候変動枠組み条約締約国会議(COP15)  
宇宙航空研究開発機構
- 五十音順(アルファベット順)にならべられた事典においては、用語の先頭の語を思い出さねば事典が引けない。

- 自分の思いついた部分的な語によって事典が引けるようにすることが必要である。
- KWICインデックス、部分マッチングによる全文検索が必要となる。

# Wikipediaにおける問題点

- 用語選択における体系性があるか。
  - 同義語の中で最も妥当なものが見出し語として選ばれているか。
  - 用語の粒度について配慮されているか。
- 数十万語から百万語といわれるWikipediaの項目をオントロジーの立場からクラスタリングし手薄な分野を見つけ補強する。

- 見出し語について網羅性が確保されているか。

学問の各専門分野では数万～数十万の用語が存在するし、専門分野の数は少なくとも数百はあると言われている。

- したがって、数百万から1千万語程度の見出し語まで努力して増やしていく必要があるだろう。
- 見出し語や用語について関連する用語へのリンクをもっと密につける。

# 用語解説についての信頼性

- Wikipediaの記述は信頼性に欠けると言われることがある。
- 多くの自由人の参加による用語解説で厳密な学問的正確性を保証することは難しい。
- 多くの事典、本、雑誌、新聞などにおける関連する記述との相互参照を行い、信頼性を確保する。

# Web情報の信頼性

- Webサイトで得られる情報の信頼性が問題となっている。
- 信頼性の推定は内容のチェックと発信者の身元情報から行われている。
- ある問い合わせに対して、どのような意見がどのように分布しているかを示すことも役に立つ。

# 辞書・事典の維持・改善

- 新しい語、新しい分野が毎月のように現れる
- これらの新しい語を辞書、事典に半自動的に取り入れることが必要となる。
- 電子新聞、電子雑誌に現れる新語の発見とともに、何カ月、何年かして消えてゆく語についても追跡をすることが大切である。
- 新しい分野の認定、そこでの用語の収集とオントロジーの作成にはその分野の専門家を必要とする。

# 用語の社会学

- 各地方の新聞やウェブのテキストを分析することによって、方言、語の用法の地域的分布、年齢分布、職業分布、時代分布等を明らかにする。
- 新しい語が作られ、それが流布し、やがて消滅してゆく間（語のライフサイクル）の諸相を調べる。
- これを言葉の生態学と名づけて、その学問的重要性を指摘した（1980年代）

# 私の作りたい辞書

- 専門用語辞典としては岩波情報科学事典(1990)を作った。これはいろいろな観点から引ける冊子体の辞書としてはどこにもない一つの完成形である。
- 国語辞書としてこれまでには作られていない“意味の逆引き辞書”を作りたい。
- もう一つ作りたい辞書は“言葉の連想辞書”である。

# 意味の逆引き辞書

- 主義・主張を変える  
変節、変心、転向、無節操、  
背信、背徳、不義、裏切り
- なすこともなく退屈  
無聊、倦怠、徒然、アンニュイ
- 物が壊れる  
毀損、破損、破壊、損壊、瓦解
- 補う、不足を埋める  
補填、補充、補遺、増補、補正
- あるものの様に似せる  
擬態、カムフラージュ、模擬  
パロディ、換骨韃鞮、剽窃、盗作
- あらわれたり消えたりする  
変幻、変幻自在、出没  
融通無碍、神出鬼没、臨機応変

# 言葉の連想辞書

- 70歳、77歳、・・の別称を思い出せない。ただ61歳が還暦ということは知っている。そこから70, 77歳の別称が出てこないか。ウィキペディアでやってみると何とか出てくるが、かなりの手間がかかる。
- 還暦                    古希、喜寿、米寿、白寿
- バリトン              テノール、バス、(女性)ソプラノ、メゾソプラノ、アルト