7E−4

二か国語辞典における辞書記述項目の設計

荻野孝野　　横山晶一
（日本電子化辞書研究所）　　（電子技術総合研究所）

１．はじめに

本報告は、Tubingen大学江沢建之助教授を中心にして進められた二か国語辞典の開発の一環に関するものである。近年の日本語学習者の増加に伴い、役に立つ日本語辞典の出現が望まれている。従来の二か国語辞典は、日本人が外国語を学習するための辞書であって、日本語側の語義や文中での文法的ふるまいを知っていることが前提となっているため、 逆に外国人が利用するには、日本語としての情報が不十分である。我々は、これらの問題を踏まえ、必要な辞書記述項目や辞書項目の配置などを検討し、試作的に記述を進めてみた。ここでは、これらの問題の中で、特に辞書記述項目の配置を中心に述べる。

２．辞書開発の概要

プロトタイプとして、 基本語約２０００語、 漢語約６００００語、専門語約１００語について記述を進めている。

記述は、日本語側を日本人が記述し、独語側を日本語あるいは英語のわかるドイツ人が記述する形で進められている。他に、文化庁の「外国人のための基本語用例辞典第二版」の形容詞・副詞・名詞に関する文例約２０００文についても、独語訳文をつけ、共同研究者である石川徹也氏（図書館情報大学）により、日独語文例検索システム[2]が開発された。

３．辞書記述項目の配置

3.1 辞書記述項目の配置の多様性

辞書記述項目の並べかたとして、語義の違いを第一キーとするか、文法的情報の違いを第一キーとするか、子見出し(親見出しを構成語の一部として含む、複合語や意味的関連のある派生語など)を関係する見出しの記述内に配置するか、見出し扱いするかなど、記述項目の配置にさまざまなバラエティがある。

例１は、初期段階の記述例であるが、これによると、訳語は元言語の品詞と同じものが挙がっているが、例文では、概念が同じ別品詞の単語が使われていたり、同じ概念の構成語を含む複合語が、所属する親見出しの語形の違いにより、別見出し部分に配置されたりしている。訳語や語義の近さからみても近い場所に配置したい。例えば、これを表現辞典として用いるならば、 例２．１に示すように語義を中心にした配置がある。検索効率を重視するならば、例２．２に示すように、子見出し親見出しの区別をつけず、見出しのコード順に配置するのがよい。こういった辞書項目の配置については、従来の辞書作成では、出力イメージに併せて企画の段階で設定しなければならなかったが、現在は、計算機環境を導入することによって、出力形態に拘束されないで辞書記述項目を配置でき、柔軟な辞書開発が可能となった。

A design for dictionary's description items in the bilingual dictionary.

Takano Ogino[1], Shoichi Yokoyama[2]
1. Japan EDR Institute Ltd., 2. Electrotechnical Laboratory

例1　初期段階の記述
##あか## ①[赤] N ROT<n>:(→あお ①[青]，まっか ③[真っ赤])
Rot<n>:信号<シンゴウ>の赤<アカ>は「止まれ」という意味です. Rot bei der Ampel bedeutet "Halt"./Wenn die Ampel rot ist, heißt das "Halt".
Rosa<n>:赤のご飯を炊いてお祝いをした. Wir haben "roten Reis"(mit Azuki gekochter Reis für festliche Anlässe) gekocht und gefiert./Wir haben rosa Reis gekocht, um zu feiern. $～いぬ$ ⓪ [～犬] N brauner Hund<m>
##あか・い##[赤い] ADJ ROT:(～さ⓪ N)
rot:夕日で西の空が赤かった. Der Himmel im Westen war rot von der Abendsonne.
$～いはね$⓪ [～羽根]rotes Federabzeichen für Spender bei öffentlichen Sammlungen im Oktober.

◇配置の多様性

例2.1
1 ##あか##[赤]
　語義1
　語義2
　語義3
2 ##あか・い##[赤い]
　語義1
　語義2

例2.2
1 ##あか##
　1.1 色が赤い.
　　1.11##あか##[赤]
　　1.12##あか・い##[赤い]
　　1.13##あか・さ##[赤さ]
　1.2 色が赤さの程度.
　　1.21##あかさ##[赤さ]

3.2 辞書記述形式

我々は、辞書記述形式の設定にあたって、

・辞書項目のそれぞれに一意的な識別記号を設定する

・機械的にできる配置とそうでないものを見極め、人手による判断の必要な部分を初期記述で行うことに留意し、記述後のデータ変換が容易になるような形式をめざした。図1に示すような構造に従い、例3のように記述できる。

図1.　記述項目の全体配置イメージ
1　識別番号
2　幹見出し情報（ひらがな見出し）
　2.1 表層表現情報
　　2.11よみ情報
　　2.12表記
　2.2 基本意味（個別意味全体の共通概念）
3.i 個別意味情報（使用頻度の高いと思われるものから記述）
　3.i1 意味情報
　　3.i11個別意味
　　3.i12関係語指示
　　3.i12m関係語/関係
　3.i2 形態・文法情報
　　3.i21語形(異表記，語尾変化)
　　3.i22品詞
　　3.i23表層的共起関係（<意味区分><格表示>の組み合わせ>
　3.i3 訳語情報
　　3.i31形態・文法情報
　　3.i311品詞
　　3.i32用例

例3 図1に基づく記述例
#あか#a┐ka[赤，赤い，赤さ，赤み，赤める，赤らむ] {ROT|色が赤いこと. }
(1) {色が赤いこと. |ROT} $真紅＋,あかね色＋,赤色＋,真っ赤だ＋$
　(1.1)##あか##a┐ka[赤]名詞　$真紅＋,あかね色＋,赤色＋$
　　#Rot#N
　　:Rot<n>:信号<シンゴウ>の赤<アカ>は「止まれ」という意味です.
　　|Rot bei der Ampel bedeutet "Halt". /Wenn die Ampel rot ist, heißt das "Halt".
　　#Rosa#N
　(1.2)##あか・い##a┐kai[赤い]形容詞
　　//<具体物>ガ～//
　　#rot#ADJ
　　:リンゴは熟すと赤くなります. |Wenn die Äpfel reifen, werden sie rot.

4．今後の課題

構造的な辞書項目の設定に基づいた記述の試行を更に前進させ、記述内容、語数を充実させ、実用的、高品質の二か国語辞典として発展させたい。

□参考文献 1 江沢他：日独語電子化辞書システムの構築，トヨタ財団1988年度第三種研究助成経過報告会資料
2 石川徹也：文例検索システム，2nd International Confernce on Japanese Information in Science ,Technology and Commerce