

『中国成語デジタル大辞典』まもなく完成

中国の成語には四字句に民族の智慧と文化が凝縮され、その簡潔な言語形式とあいまって、古代から現代まで様々なジャンルで頻繁に使用される。日本語にも中国語由来の成語が多数あり、標語やキャッチコピーに至るまで広く愛される表現形式のひとつである。

北京大学計算言語学研究所 (Peking University Institute of Computer Languages : PICL と略称) が中国国家 973 プロジェクトで開発を進めて来た『中国成語デジタル大辞典』(中文名《成語知识库》) がまもなく完成する。本デジタル辞典は、PICL の現代中国語語法情報コーパス (中文名《现代汉语语法信息词典》) の成語 5000 条を核に【表 1】、自然言語処理理論と専門技術に依って計算した使用頻度・分布密度情報に基づき、「人民日報」ならびに『成語大全』『学生成語詞典』など中国出版の代表的成語辞書 5 冊から 3 万項目余の現代中国語成語を収集してデータベース化し語法属性や英訳などを付すものである。

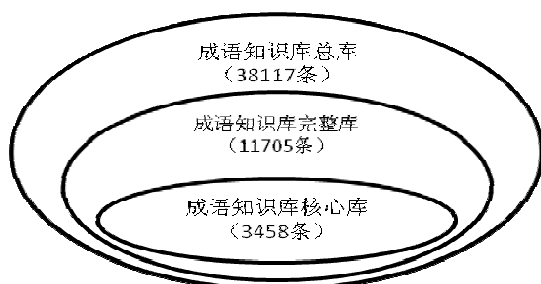


表 1 中国成語デジタル大辞典の組織図

周知のように中国語成語は、典故の由来、表す意味などに加え、その語構成や品詞機能が分かると成語の使用に便利である。『中国成語デジタル大辞典』の注釈は PICL 現代中国語語法情報コーパスから受け継いだ 16 種 (追加予定を含む)

の言語属性が付されている。

NO、難度レベル、句構成【表 3】、品詞機能【表 4】、ピンイン、異形、類語、反義語、直訳、意識、英訳、英訳来源、釈義、頻度、典故、動画 (追加予定)、語感情報 (追加予定)

表 2 言語属性一覧

属性	句構成	成語例
pz	偏正关系	孜孜不倦、一衣带水
zw	主谓关系	草木皆兵、枯木逢春
bl	并列关系	不学无术、千锤百炼
db	动宾关系	蹉跎岁月、饱经风霜
dbu	动补关系	重于泰山、无动于衷
dbb	动宾补关系	问道于盲、入木三分
ld	连动关系	打草惊蛇、兴妖作怪
jy	兼语关系	请君入瓮、唆人作恶
fz	复指关系	半斤八两、半夜三更

表 3 句構成一覧と実例

IN	名詞機能成語
IB	区別詞機能成語
IA	形容詞機能成語
IM	修飾機能成語
ID	副詞機能成語
IV	動詞機能成語
IC	補語機能成語

表 4 品詞機能一覧

この基本データを使用すれば、成語関連の出版や翻訳、中国語教育の科学性と編集効率向上に寄与するに違いない。日本の成語との対照や、日本文学作品中の中国成語の使用状況などの情報を加え、日中比較文化研究での応用も可能だ。

本辞書は蔣經國國際學術交流基金會による国際プロジェクト「歴代語言知識庫の構築」(研究代表台湾元智大學中國語文學系講師羅鳳珠)の一環として構築中される <http://cls.hs.yzu.edu.tw/DLKB/Index.aspx> 上掲資料は北京大学計算言語学研究所王雷、常宝宝先生提供によっている。

文責：砂岡和子 早稲田大学政治経済学術院