

MORPHOLOGICAL CONSTRUCTIONS AND DERIVATION OF
SENSES AND PART-OF-SPEECHES FOR CHINESE COMPOUNDS

You-Shan Chung * **Keh-Jiann Chen**

Academia Sinica, Taiwan

ABSTRACT

Since new compounds are generated very productively in Chinese, an automatic scheme is required to predict their part-of-speech and senses in order to automate computer language processing. To this end, we analyzed the morpho-syntactic behaviors of about 4,025 most productive morpheme characters in our affix database. We found that semantic and logical compatibility are more important than syntactic constraints in compounding. Hence, we classified morphemes into four major semantic types: object, act, attribute and value, and use semantic composition rules to derive the meaning and part-of-speech of compounds. Although some morpheme types and composition rules are ambiguous, we propose resolutions like constraints, analogy, morpheme position, and morpheme-specific rules to deal with them.

KEYWORDS

Chinese compounds Semantic type Affix Part-of-speech prediction
Sense disambiguation

中文複合詞的構詞及其語意和詞性衍生規律

鍾友珊 陳克健

中央研究院，台灣

提要

中文複合詞的衍生性很強，需要一套自動預測詞性和語意的方法，以便於機器處理。因此，我們挑選了約四千零二十五個衍生性最強的詞素，分析他們的構詞與語法行為。因為我們發現就複合詞的形成來說，語意和邏輯性的影響大於語法的影響，於是將這些詞素分成四個主要語意類別：物體、動作、屬性、屬性特徵，並針對不同類別的組合訂定語意合成規則，得出各種組合的語意及詞性。由於有些詞素本身和語意合成規則有歧義，我們也提出了一些輔助性的判斷標準，例如類比、位於詞首/詞尾，以及針對部分詞素個別訂定的規則。

關鍵詞

中文複合詞 語意類別 詞綴 詞性預測 語意解歧