

A lexicon of multiword expressions for linguistically precise, wide-coverage  
natural language processing

田辺 利文\*      高橋 雅仁†      首藤 公昭‡  
Toshifumi Tanabe    Masahito Takahashi    Kosho Shudo

出典 : Computer Speech & Language, Volume 28, Issue 6 (Nov. 2014), pp. 1317–1339

Since Sag et al. (2002) highlighted a key problem that had been underappreciated in the past in natural language processing (NLP), namely idiosyncratic multiword expressions (MWEs) such as idioms, quasi-idioms, clichés, quasi-clichés, institutionalized phrases, proverbs and old sayings, and how to deal with them, many attempts have been made to extract these expressions from corpora and construct a lexicon of them. However, no extensive, reliable solution has yet been realized. This paper presents an overview of a comprehensive lexicon of Japanese multiword expressions (Japanese MWE Lexicon: JMWEL), which has been compiled in order to realize linguistically precise and wide-coverage natural Japanese processing systems.

---

\* 福岡大学

† 久留米工業大学

‡ 福岡大学名誉教授