

通時コーパスの設計—近代語コーパスの展開と活用—

田中牧郎 (言語資源研究系)

■太陽コーパス

- 明治後期～大正期の総合雑誌『太陽』から5年分を抽出した全文コーパス
- 構造化テキストタグ付きコーパス
- 1895(明治28)年、1901(明治34)年、1909(明治42)年、1917(大正6)年、1925(大正14)年の、各年12冊、計60冊
- 総文字数: 約1450万字
- 著者数: 約1000人
- 2005年公開

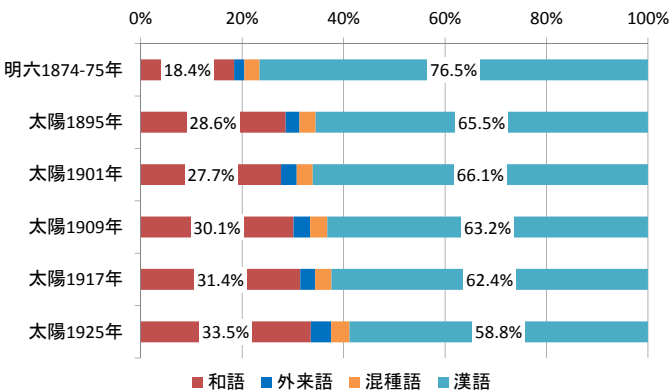
■近代女性雑誌コーパス

- 明治後期～大正期の女性雑誌3種から40冊を抽出した全文コーパス
- 構造化テキストタグ付きコーパス
- 『女学雑誌』1894(明治27)年・1895(明治28)年、『女学世界』1909(明治42)年、『婦人倶楽部』1925(大正14)年
- 総文字数: 約210万字
- 著者数: 約480人
- 2006年公開

■明六雑誌コーパス

- 明治初期の学術啓蒙雑誌『明六雑誌』(1874～1875年)全文コーパス
- 形態論情報(単語情報)付きコーパス
- 原文画像参照機能付き
- 総語数: 約18万語
- 著者数: 16人
- 2012年公開

■語種構成の変化



■語彙レベルの変化 (太陽コーパス(1895～1925年))

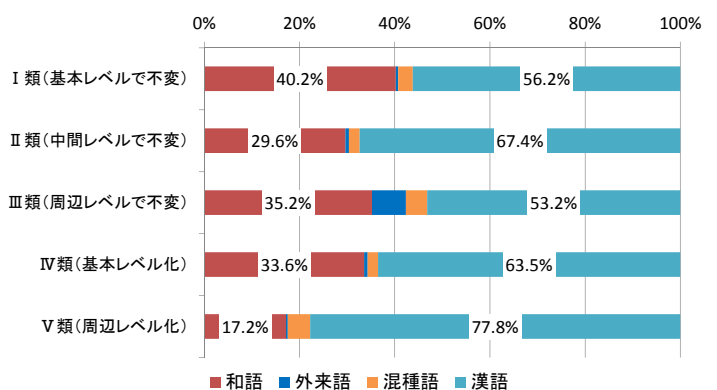


図1 『明六雑誌コーパス』『太陽コーパス』の語種構成の変化

図2 『太陽コーパス』の語彙レベルの推移

基本レベル=基本語彙
周辺レベル=周辺語彙

■漢語「実現」「表現」の基本語化とその要因

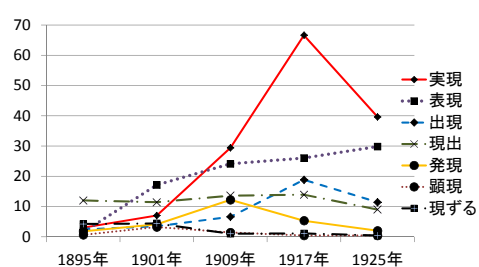


図3 「あらわす」「あらわれる」の類義漢語の出現率

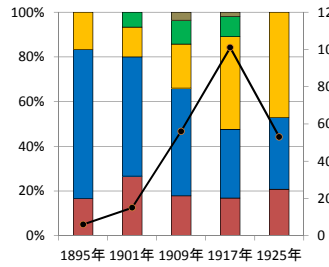


図4 「実現」(他動詞)の対象と総頻度

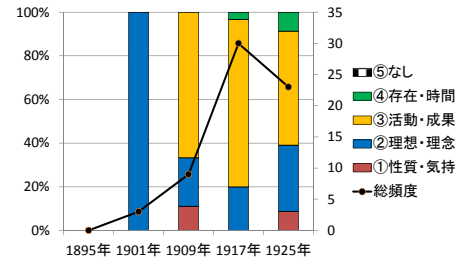


図5 「実現」(自動詞)の主体と総頻度

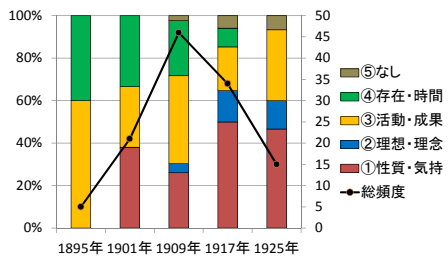


図6 「表現」(他動詞)の対象と総頻度

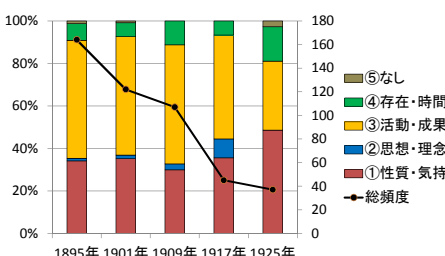


図7 「あらわす」の対象と総頻度

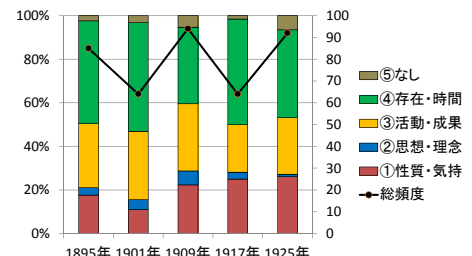


図8 「あらわれる」の主体と総頻度

表1 「あらわれる」と「実現」(自動詞)の主体の語句(「③活動・成果」の場合)

	「あらわれる」の主体	「実現」の主体
活動にあたる語句	事実、記事、宣言、話、作品、文、歌、文字、現象、現実、問題、出来事、事、行動、発達、変革、改革、気風、習慣、制度、(ほか多数)	統一、合同、独立、平和、政治、改正、引退、更迭、協力、借款、撤廃、注文、選挙、利下げ、物価を下落せしめること、(ほか多数)
成果にあたる語句	結果、影響、成績、功績、名、實、弊害、弊、(ほか多数)	結果

①内面にある性質や気持ち

さう云ふ志を實現するには (1917年6号・副島義一「選挙の実験」)

②目指される理想や理念

是の大理想を實現するに先ちて (1901年13号・高山樗牛「文芸時評」)

③社会的な活動や成果

人類の平和安穩なる生活を實現することを得べし (1909年13号・上杉慎吉「我が立憲政体と大政党」)

④具体的な存在や時空間

黄金國を實現することが出来るといふ程でなくとも (1909年4号・近松秋江「文士の見たる政治家」)