

言語資源研究系
独創・発展型共同研究プロジェクト

「近代語コーパス設計のための文献言語研究」

リーダー： 准教授 田中牧郎

国立国語研究所のコーパス（書き言葉）

800 1000 1200 1400 1600 1800 1900 2000年

古典語

近代語

現代語

「通時コーパスの設計」プロジェクト
で設計中

1975～2005年（構築済み）
「現代日本語書き言葉均衡コーパス」

1868～1945年 本プロジェクトで設計（共同研究員：10名）

- ・近代語を代表できるコーパスを設計する
- ・古典語のコーパスと現代語のコーパスを接続する
- ・コーパスの作成技術や利用ツールは古典語・現代語と共有
- ・文献と言語が多種多様な近代に特有の課題を重点的に研究

コーパス近代
語研究の開拓

- ・文法研究
- ・語彙研究
- ・表記研究
- ・文体研究

文献研究と
文献選定

- ・文献リストの
作成と分類
- ・資料選定基準
の研究

文献言語の構造
化と形態素解析

- ・構造化タグの
設計
- ・形態素解析
辞書の整備

学会等を通して
成果を公表

「近代語コーパス」の設計書を作成し、これを用いて次期プロジェクトでコーパス構築に着手