国会会議録に基づく自民系/民主系大臣の特徴分析

樋口心 掛谷英紀

[S1511175@u.tsukuba.ac.jp](mailto:S1511175@u.tsukuba.ac.jp) [kake@iit.tsukuba.ac.jp](mailto:kake@iit.tsukuba.ac.jp)

筑波大学

**概要**本論文ではWeb上で公開されている国会会議録を用いて政治家として正しい議論を行えない政党/人物の特徴を見出すことを目的とする。昨今野党により問題になっている国会答弁内の「ヤジ」について注目し、民主党政権時代の野党と現在の野党の発言の分類を試みる。

具体的にはナイーブベイズ法を用いた機械学習、また両カテゴリの使用頻度の偏りから特徴語の抽出を行う。

使用頻度の偏りから、現在の野党は自分のマニフェストではなく相手の言質を取るような発言が多いという傾向が分かった。また、末尾表現に着目した機械学習の分類は高い精度を誇り、政治テーマに依らずその個人/政党による特徴を抽出することに有効であることが示された

**２．研究手法**

国会会議録を対象とした自然言語処理に関する研究はいくつかあるが、本研究では末尾表現に着目し各政党の言葉遣いを判別し、有意義な議論を行う党を判別するシステムの構築を行う。

　まず、国会会議録検索システムから国会会議録の答弁データを入手し、形態素に分割する。

その後、読点の直前3語または４語を末尾表現と定義し、表出形をそのまま機械学習の素性として使用する。

また、動詞・形容詞・名詞なども追実験の対象とした

**3．取得データ**

**3.1国務大臣**

民主系野党と自民系野党を比較するために、

民主系の野党は2009/9/16~2012/12/25,

自民系の野党は2015/4/16~2018/7/31

の3年間の言語データを収集した。上記の期間に発言を行った、つまり言語データに含まれる大臣のリストを以下に示す。

**3.2国会会議録**

本研究では言語データを収集するために、国会会議録検索システム検索用APIを利用して分析対象の大臣の発言を収集する。

先述の通り、民主系および自民系の過去の国会における発言から各カテゴリの特徴分析を試みる。

（収集元定義）

**4．実験**

**4.1データセット**

3.2節で収集した発言を表

**4.2機械学習**

上記のデータセットに基づき、ナイーブベイズ法による機械学習を行った結果を表4に示す。

**4.3単語出現頻度**