

8. 政治テキストの計量分析

石田基広・金明哲編『コーパスとテキストマイニング』共立出版、2012年、97-106頁

大賀哲（法学研究院）toga@law.kyushu-u.ac.jp

- ・ 政治テキストとは → 政治現象に関連するテキスト（新聞、雑誌、演説）
- ・ データベースの整備と言語処理技術の発達
- ・ 学問分野の枠にとどまらない学際的な研究

※参考文献等は本書参照のこと（代表的なもののみ例示）

8.1 政治テキスト分析の概観

- ・ 「内容分析」としての政治テキスト分析：Lasswell (1948)
→ マスメディア研究 プロパガンダ研究
- ・ 内容分析の定義：「テキストにおける特定の特徴を体系的かつ客観的に明らかにすることで、推論を行う研究手法」（Stone *et al.*, 1966）
- ・ 内容分析＝音声／映像／テキスト分析 但し政治学ではテキスト分析が主
- ・ マスメディア研究や政治学の研究→新聞演説などの単語やカテゴリーの研究
- ・ 総理大臣国会演説の重要語彙調査（田中 1994、1999、Tanaka 1994）
- ・ テキストを分析することによって政治現象の分析に何らかの知見を得ること
 - ① 「何がテキストにおいて述べられているか」
 - ② 特定の内容語やカテゴリーの頻度
 - ③ 仮説—演繹的分析
- ・ 政治テキストの言語学的研究（述べ語数、異なり語数、文長、単語長、品詞比率）
- ・ 演説文（大統領演説、総理大臣演説）の特徴分析や国会議事録データベース
- ・ 言語学研究の特徴→テキストそれ自体の特徴
 - 「どのようにテキストにおいて述べられているか」

8.2 政治テキスト分析の実際

- ・ 総理大臣国会演説→日本政府の政策や認識を表す素材
（通時的な分析に耐えうる公開資料）
- ・ データベース「世界と日本」（東京大学東洋文化研究所）
<http://www.ioc.u-tokyo.ac.jp/~worldjpn/>

8.2.1 語彙多様性に関する研究

- ・ 施政方針演説と所信表明演説テキストの特徴量／政治的役割の差異

- ・ 中曽根・小泉の演説テキストの特徴量／政治的役割の変化

①名詞についての述べ語数 N (出現する語の総数)

②述べ語数・異なり語数比 TTR → (出現する語の種類) 語彙多様性の指標

③シン普森の D (相対的集中度) → 特定の語の反復率

- 所信表明演説よりも施政方針演説のほうが N の値が高い
- 施政方針演説と所信表明演説では TTR 、 D の値に有意な差は見られない
 - ・ 施政方針演説 (通常国会) / 所信表明演説 (特別国会 / 臨時国会 / 首相交代時)
→ 前者がより包括的な内容 / 後者がより選択的内容
 - ・ レジスター制約の強さ
- 鈴木以前よりも中曽根以降の方が N の値が高い
- 小泉において TTR の値が有意に高く、 D の値が有意に低い
 - ・ 国会演説量の増加
 - ・ 多様な語彙を盛り込みつつ、特定の語彙が集中して用いられているわけではない

8.2.2. 対外認識に関する研究

個別の内容語の研究 / 構文全体の分析 → 日本外交の大局的推移

① 国名・地域名の抽出とカテゴリー化

② 「日米」を含む複合名詞を抽出し、安全保障 / 経済 / 文化で分類

③ 「日米」を含む文節に係る文節、「日米」を含む文節に係る文節

→ 冷戦以前 / 以後の特徴量

- 第一主成分は欧米地域の絶対値が高く、第二主成分はアジア地域の絶対値が高い、第三主成分はアジアが比較的高い
- 第一主成分は冷戦崩壊後減少、第二主成分は70年代以降増加
- 第一主成分は対欧米外交の軸、第二主成分が対アジア外交の軸、第三主成分が危機対処の軸

- 冷戦崩壊後、日米関係の言及は減っているが、安全保障の言及割合は増えている。
- 冷戦崩壊後、日米関係は「国際協調」と併置される
- (係り受け) 冷戦崩壊以前は日米 → 友好、冷戦崩壊後は日米 → 「信頼」「機能」「連携」「協力」(文脈は安全保障)

8.3 おわりに

- ・ データベースの整備と言語処理技術の発達 → 学際連携が容易
- ・ 政治学と言語学の融合研究としての「政治テキスト分析」