

地位と感情表出

——国会会議録と感情辞書を用いた分析——

関西学院大学 中野康人

【1. 目的】

本報告の目的は、国会議事録をデータとして用い、発話者の社会的地位と発言内容から抽出される感情との関係を記述的に探索することにある。テキストデータを情報源とする分析の手法や技術は近年大きく発展しており、社会学的な分析の対象として、蓄積されたテキストデータを利用する研究事例も広がっている。しかしながら、記述内容つまりはそこに含まれる意味・意識・感情などの測定は依然として質的な解釈に大きく依存している。この報告では、テキストに表象される発話者の感情を量的に集計し、時系列的变化や討議内容・発話者の属性などとの関係を紹介する。

【2. 方法】

標準的な調査票による社会調査が困難になる一方で、テキストデータを情報源とする分析の手法や技術は近年大きく発展している。報告者も、これまで調査票調査の自由記述や新聞記事といったテキストデータの分析に取り組んできた（中野 2010 など）。本報告で分析の対象となるのは、国会の会議録である。日本における国会の会議録は、1947年の第一回から今日に至るまで、国立国会図書館の検索システムで公開されている (<http://kokkai.ndl.go.jp>)。発話者の発言そのままの記録ではないが、70年にわたる国政の話題が、時間や発話者の記録とともに蓄積・公開されている。

【3. 結果】

今回の報告では、複数の発言者について、立法府・行政府における地位、年齢、発言場所、発言内容を整理したデータを準備する。その上で、テキストデータから、その書き手や内容の発話者が持つ感情や意識を測定する。分析に使用する辞書は、感情極性評価辞書（小林ら 2005）と NRC Word-Emotion Association Lexicon である。個人の効果、年齢の効果、年代の効果、話題の効果などの識別に注意しながら、地位と発言内容の感情極性の関係をまとめる。

【4. 結論】

分析の結果、個人の効果だけでなく、地位による感情極性の違いが認められた。本報告では、複数の感情辞書による分析結果を比較し、分析の安定性を確かめる。また、数十年の過去に遡る通時的なデータを分析する際の課題、共通の内容を国際比較する際の課題を提示し、発展的利用の方向性を探る。

【文献】

小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一, 2005, 「意見抽出のための評価表現の収集」, 『自然言語処理』, 12(3):203-222.

Mohammad, Saif., 2011, NRC Word-Emotion Association Lexicon (<http://www.purl.org/net/NRCemotionlexicon>).

中野康人, 2010, 「読者投稿の記述的計量テキスト分析 --- 「声」と「気流」 ---」, 『関西学院大学先端社会研究所紀要』, 2:43-57.