昨今、野党の国会に対する姿勢は疑問視される。2017年の後半から現在にかけて予算委員会の討論では森友・加計問題が主なテーマとして話し合われてきた。半年以上確固たるデータが得られないなか、安部首相が関与していたかどうかを野党は激しく追及しており、実際に国会の質問時間の配分は「与党2割・野党8割」という状況に陥っている。国会では本当に話し合うべき話題が尽きないにも関わらず、与党の失言を取ることに会議の大半の時間を使っている現状は改善されるべきである。また、私たち国民も適任と思える立候補者に投票し、国会を作っていく必要がある。しかし、メディアによる報道では恣意的な印象操作が行われている可能性もあり真相がわからない。

　そこで、私はWeb上にあるデータに自然言語処理を行い、現在の野党の発言は適切かどうかを定量的に調べる。投票するべき人物を選ぶ指標になる研究結果を打ち出し、国会の現状を是正したいと考える。現在行っている研究では国会会議録APIから議員の予算委員会内の発言を収集し、野党の言葉遣いについて調べている。具体的な手順は以下の通りである。

まず民主党政権時代の野党と、現在の野党の発言データを二つのカテゴリーとして年代・人物で分けて収集する。そして集めたデータを形態素解析し、動詞・形容詞・末尾表現等、会議のトピックに寄らない成分のみを取り出す。そして、二つのカテゴリーの中で出現頻度が偏っている言葉を特長語としてとらえる。また、機械学習アルゴリズムにかけ、どの議員が各カテゴリーに分類されるか、カテゴリーを特長づける単語は何か等を調べる。この研究結果をもとにどういった人物・派閥が会議を冗長にしているか、国民・国家のために必要な議題を話しているか等を判別できる。

　今後の展望としては、収集するデータの拡大、データの活用方法の変更の2点を図る。国会会議録APIによる議事録データのみに限らず、Twitterの発言や著書など多角的な文章から分析を行い、さらなる特徴をつかむ。また、政党ごとの特徴を収集したうえで、特徴による議員・政党のマップ化を行い、視覚的に多くの人に訴えられるような研究内容にしていくことを目指す。