

Measuring Changes in the Collective Identity of Political Elites in Singapore (1955—2011)
using Text Analysis on a Large Corpus of Parliamentary Discourse

by

Gregory Dalziel
Keio University
Graduate School of Media and Governance

Abstract

Does collective identity change over time? And how can we measure such changes? This dissertation focuses on three inter-related questions: (1) how groups perceive themselves; (2) how they perceive others; and (3) the perceived relationship between themselves and others. Most literature in international studies sees identity as fixed and measuring identity is often carried out on small samples that are difficult to replicate. The availability of 'big data' and computational methods provides an opportunity for empirically measuring identity and identity changes. Our study is located in Singapore; its history of merger with Malaysia and subsequent independence offers a useful site to study identity changes. There is little research that systematically measures elite identity in Singapore.

For this study we constructed a unique dataset: the Singapore Parliamentary Speech Corpus, containing almost 41 million words. It is the largest English-language dataset outside of Europe and North America of its kind.

To answer each of our three questions we use text mining methods to measure the vocabulary of identity and use correspondence analysis to construct an identity space. This enables us to measure not only changes in identity vocabulary but also changes in time.

Our findings indicate that elite self-understandings changed dramatically before and after independence. Who they paid attention to also changed from a regional focus to a more international one over time. The perceived relationship between Singapore and others can be generally divided into: (1) economic competition; (2) emulation and the borrowing of policies; and (3) an overarching category of comparison. We also carry out an in-depth case study of comparison, extracting almost 1,500 direct social comparisons and find that comparisons changed from one of competition and status attainment in the 1980s to one of monitoring rankings and image management in the 1990s and 2000s.

議会演説のテキスト分析によるシンガポールの政治エリートの集団的アイデンティティ変化の計測（1955～2011年）

ディーエル グレゴリー

（内容の要旨）

集団的なアイデンティティは時間とともに変化するのか。どのようにすればそうした変化を計測できるのか。本研究は、三つの相互に関連する問題に焦点を絞る。つまり、第一に、どのように集団が自身を認知するのかということ、第二に、どのように彼らが他者を認知するのかということ、第三に、自己と他者との間の認知上の関係である。国際関係研究のほとんどの文献はアイデンティティを所与のものとして見ており、アイデンティティの計測は、再現が困難な小さなサンプルに基づいて行われることが多い。しかし、「ビッグ・データ」とコンピュータの利用によって、アイデンティティとその変化を実証的に計測する機会が提供されるようになってきている。本研究が扱うのはシンガポールである。マレーシアとの合併とそれに続く独立の歴史は、アイデンティティの変化を研究する場所として有用である。シンガポールにおけるエリートのアイデンティティをシステムティックに計測した研究はほとんどない。

本研究のために独自のデータセットを構築した。シンガポール議会演説コーパスは、約4100万語からなる。これは非欧米圏で最大の英語データセットである。

上述の三つの問題に答えるため、テキスト・マイニング手法を用いてアイデンティティの語彙を計測し、コレスポネンス（対応）分析を用いて「アイデンティティ空間」を構築した。これによってアイデンティティの語彙における変化を計測するだけでなく、時系列変化を計測することもできる。

分析の結果、エリートたちの自己認識は独立の前後で劇的に変化していることが分かった。誰に注目を払っているかも、地域的なフォーカスから、よりいっそう国際的なものへと、時間が経つにつれて変わっている。シンガポールと他者との間の認知上の関係は、徐々に三つに分かれていった。第一に、経済的な競争、第二に、政策の模倣と借用、第三に、競争に関する包括的なカテゴリーである。自己と他者の比較についてのケース・スタディを行い、約1500の社会的な比較についての文章を抽出した結果、1980年代の競争とステータスの獲得から、1990年代および2000年代のランキング重視とイメージ管理へと変化していることが分かった。